

RUNNING HEAD: Planning to speak in L1 and L2

Planning to Speak in L1 and L2

Agnieszka E. Konopka^{1,2}

Antje Meyer²

Tess A. Forest^{2,3}

¹University of Aberdeen
Scotland, UK

²Max Planck Institute for Psycholinguistics
Nijmegen, The Netherlands

³University of Toronto
Canada

Address correspondence to:
Agnieszka E. Konopka
University of Aberdeen
United Kingdom
agnieszka.konopka@abdn.ac.uk

Abstract

The leading theories of sentence planning – *Hierarchical Incrementality* and *Linear Incrementality* – differ in their assumptions about the coordination of processes that map preverbal information onto language. Previous studies showed that, in native (L1) speakers, this coordination can vary with the ease of executing the message-level and sentence-level processes necessary to plan and produce an utterance. We report the first series of experiments to systematically examine how linguistic experience influences sentence planning in native (L1) speakers (i.e., speakers with life-long experience using the target language) and non-native (L2) speakers (i.e., speakers with less experience using the target language). In all experiments, speakers spontaneously generated one-sentence descriptions of simple events in Dutch (L1) and English (L2). Analyses of eye-movements across early and late time windows (pre- and post-400 ms) compared the extent of early message-level encoding and the onset of linguistic encoding. In Experiment 1, speakers were more likely to engage in extensive message-level encoding and to delay sentence-level encoding when using their L2. Experiments 2-4 selectively facilitated encoding of the preverbal message, encoding of the agent character (i.e., the first content word in active sentences), and encoding of the sentence verb (i.e., the second content word in active sentences) respectively. Experiment 2 showed that there is no delay in the onset of L2 linguistic encoding when speakers are familiar with the events. Experiments 3 and 4 showed that the delay in the onset of L2 linguistic encoding is not due to speakers delaying encoding of the agent, but due to a preference to encode information needed to select a suitable verb early in the formulation process. Overall, speakers prefer to temporally separate message-level from sentence-level encoding and to prioritize encoding of relational information when planning L2 sentences, consistent with *Hierarchical Incrementality*.

Keywords: incrementality; time-course of sentence formulation; L2 sentence production; linguistic experience

Producing an utterance in any language involves a complex sequence of processes, beginning with the generation of a preverbal message (message-level encoding) and proceeding with linguistic encoding to express this message verbally (sentence-level encoding). In everyday communication, speakers must be able to deploy message-level and sentence-level encoding processes quickly and fluently to generate novel utterances on the fly. As most language learners know, this goal is harder to achieve in a second language (L2) than in a first language (L1)¹. L2 production is generally slower and more disfluent, as speakers have poorer proficiency in L2 than L1: they may have smaller L2 lexicons, less familiarity with L2 syntactic structures, and crucially, less experience accessing and using L2 linguistic knowledge for communication. This latter point has received the least attention in the field, despite it being the overarching goal of any language instruction program. In fact, while there is continued interest in the L2 processing of individual words and structures as well as in control mechanisms (e.g., see [Declerck & Philipp, 2015](#); [Green, 1998](#); [Kroll & de Groot, 2009](#)), much less is known about the way that speakers *plan* L2 utterances, i.e., the way they *coordinate* message-level and sentence-level processes to produce multi-word utterances in L2.

Here we report the first series of experiments testing how linguistic experience influences processes at the interface of “thinking” and “speaking” ([Bock & Ferreira, 2014](#); [Konopka & Brown-Schmidt, 2014](#); [Konopka & Meyer, 2014](#)) in L1 and L2. We first review recent findings on incremental sentence planning and on experience-driven changes in L1 planning. Then, we report four experiments comparing the time-course of sentence planning in L1 and L2. In all experiments, native speakers of Dutch fluent in English described pictures of simple transitive events in their L1 (Dutch) and L2 (English). Their eye movements were recorded to track the way they generated preverbal messages and mapped this information onto language in real time. Thus, across experiments, we identify some of the processes responsible for L1-L2 differences in sentence planning, and we discuss the boundaries of flexibility in sentence planning.

Incrementality in L1

Producing an utterance, irrespective of the target language, involves two main classes of processes: *message generation* and then *linguistic encoding* ([Levelt, 1989](#)). *Message generation* is a pre-linguistic process that is defined as encoding of the gist of what speakers want to communicate. For example, when describing an event in which one character is acting on another character (e.g., a dog chasing a mailman; [Griffin & Bock, 2000](#)), the speaker must first encode

¹ For present purposes, we make no distinction between a *first* language and a *dominant* language. All participants in our studies used Dutch as their first and dominant language, and English as a second language.

who did what to whom in the event. This gist-based, conceptual representation includes information about the identity of the two event characters and about the relationship between them, but no lexical or structural information². The next step in the planning process is *linguistic encoding*, i.e., the retrieval of lexical items (“dog”, “chase”, “mailman”) and the assembly of a structure (an active, passive, or intransitive sentence frame) that allows the speaker to express the preverbal message linguistically.

There is general agreement that the planning process unfolds *incrementally*, i.e., that speakers generate a series of small increments of conceptual and linguistic information, rather than larger (e.g., sentence-like) units at the message level and sentence level before initiating overt production. However, there is much less consensus about the relative timing and coordination of these processes (Bock & Ferreira, 2014). Previous research in this domain showed that speakers can use different strategies when preparing simple utterances, revealing a high degree of flexibility in the time-course of sentence planning (Ferreira & Swets, 2002; Konopka, 2013; Konopka & Brown-Schmidt, 2014; Allum & Wheeldon, 2007, 2009). In particular, existing theories of sentence planning debate the *linearity* of the planning process.

One influential account (*Linear Incrementality*) posits that planning proceeds roughly concept by concept at the message level and word by word at the sentence level. For example, message-level planning can begin with the encoding of a single concept (e.g., *dog* in the current example), and encoding of a single content word (the lexical item “dog”; Gleitman, January, Nappa, & Trueswell, 2007). At the same time that linguistic encoding of “dog” begins, the speaker can start conceptual and linguistic encoding of the next message element (i.e., *chasing*), and so on until a full utterance is produced. *Linear Incrementality* thus assumes that each element of the message and sentence can be encoded independently of the next element, and that each subsequent element can be planned while speaking is already underway. As a result, message-level and sentence-level encoding are constantly interleaved.

In contrast, the second account (*Hierarchical Incrementality*) posits a smaller degree of conceptual and linguistic interleaving, and assumes instead that speakers first encode a larger preverbal message and then begin linguistic encoding (Bock et al., 2004; Griffin & Bock, 2000). Thus, on this account, speakers first generate a conceptual representation of the *who-did-what-to-whom* relational content of the message (e.g., the fact that a dog is chasing a mailman) and

² The details of this representation may vary across languages in accordance with Slobin’s (1996) “thinking for speaking” hypothesis, but we limit our discussion to messages that can be expressed in Dutch and English without large differences in structures or lexicalization preferences (see von Stutterheim et al., 2012, 2013, for a discussion of language contrasts that take such differences into account).

then begin linguistic encoding of individual message elements in the order specified by this representation (“dog”, “chasing”, “mailman”).

Evidence for both accounts comes from eye-tracking studies of sentence production. A robust finding in this literature is that, when speakers are asked to describe a visual display, they will look at the referents in this display in the order of mention and that the distribution and length of fixations to these referents will index differences in the time-course of conceptual and linguistic encoding across sentences (Bock & Ferreira, 2014; Griffin 2001; Griffin & Bock, 2000; Gleitman et al 2007; Konopka & Meyer, 2014; Kuchinsky & Bock, 2010; Meyer, Sleiderink, & Levelt, 1998). In studies eliciting transitive sentences (actives and passives), time-course analyses focus on early eye movements (0-400 ms or 0-600 ms after picture onset) and eye movements observed in later time windows (between 400-600 ms and speech onset). Linear and Hierarchical incrementality make different predictions about the distribution of speakers’ attention across a visual display in these time windows.

The strong version of *Linear Incrementality* predicts that speakers can begin formulation by encoding only one message element immediately after picture onset. Indeed, under some conditions, speakers do fixate the character that will be mentioned first very quickly (within 200 ms of picture onset) and maintain fixations on this character until approximately speech onset (Gleitman et al., 2007; Kuchinsky & Bock, 2010). The length of fixations to this character reflects the ease of name retrieval, and thus this pattern of eye movements reflects priority encoding of conceptual and linguistic information for that character alone. Conceptual and linguistic information about the second character as well as its relationship to the first character is then encoded in a separate increment, as soon as speakers finish encoding the first character.

In contrast, *Hierarchical Incrementality* predicts a clearer temporal separation between message-level and sentence-level encoding during early formulation. The hypothesis that speakers encode a larger message before beginning linguistic encoding of any part of the message is more compatible with a pattern of eye movements where speakers do not show a strong preference for one event character over another, but instead distribute their gaze *between* two characters in the first 400 ms of picture viewing. Convergence of fixations to two characters is interpreted as generation of a larger message-level representation – i.e., a representation of the event that contains more information than the identity of a single character (Griffin & Bock, 2000). Linguistic encoding is then hypothesized to unfold *after* 400 ms, as this is when speakers begin to show large, systematic shifts of gaze to the two characters in the order of mention: speakers direct their attention and maintain fixations on the first-mentioned character (the sentence subject) approximately until speech onset and then they shift their gaze to the second

character (the sentence object). The length of fixations to each character after 400 ms reflects the ease of name retrieval.

Importantly, while Linear and Hierarchical Incrementality make different assumptions about the way preverbal information is mapped onto language, they are not mutually exclusive. In fact, production studies have shown substantial variability in planning both between and within speakers. Some of this variability in planning strategies can be attributed to fluctuations in the availability of cognitive resources and extra-linguistic pressures (e.g., Ferreira & Swets, 2002; Wagner, Jescheniak, & Schriefers, 2010), but to a large extent, planning can also be controlled by processes that are specific to production itself. For example, planning can be subject to “hard-wired” grammatical constraints such as word order or morphosyntactic requirements of the target language (Hwang & Kaiser, 2014; Norcliffe, Konopka, Brown, & Levinson, 2015; Myachykov, Thomposon, Scheepers, & Garrod, 2011; Sauppe, Norcliffe, Konopka, van Valin, & Levinson, 2013). Within a language, planning is also sensitive to “softer” constraints, such as changes in the ease of completing linguistic encoding processes due to recent exposure to specific words and structures (Konopka, 2012; Konopka & Meyer, 2014; Konopka & Kuchinsky, 2015).

The key linguistic processes involved in the mapping of preverbal information onto language are lexical retrieval and structure building. The ease of retrieving words and building structures can naturally vary from sentence to sentence (Bock, 1986a, 1986b; Chang et al., 2006; Pickering & Ferreira, 2008; Smith & Wheeldon, 2001), and speakers are highly sensitive to moment-to-moment changes in the ease of completing these encoding operations: they tend to select higher-frequency words and structures more often and they produce sentences with these words and structures more quickly than sentences with lower-frequency words and structures. Importantly, these changes have implications for planning. In L1, facilitating retrieval of a character name (e.g., “dog”) results in a shift towards *linearly incremental* planning: speakers are more likely to prioritize encoding of this character during early formulation when it is easier to name. In contrast, recent exposure to a particular syntactic structure (e.g., an active structure) results in a shift towards *hierarchically incremental* planning: speakers are less likely to prioritize encoding of the subject character during early formulation of sentences using the recently experienced structure and prefer to allocate some attention to the other character in the event as well. This suggests that fast structural assembly supports generation of a larger message and thus a larger relational framework for the sentence (Konopka, 2012; Konopka & Kuchinsky, 2015; Konopka & Meyer, 2014). Such shifts from one planning strategy to another occur quickly

and, most likely, implicitly (i.e., without conscious awareness) as a result of changes in the efficiency of completing lexical and structural encoding operations.

By comparison, how does *L2* sentence planning unfold? At the moment, there is no unified framework that explains whether and how speakers might coordinate message-level and sentence-level processes differently when using their native language vs. a non-native language. Here, we apply theories that emphasize experience-driven changes in formulation (e.g., [Konopka & Meyer, 2014](#)) to make predictions about L2 planning. In general, L2 production is assumed to be slower and harder because linguistic encoding in L2 is slower and harder. As fluctuations in the ease of linguistic encoding can modulate the degree to which speakers interleave message-level and sentence-level processes in a highly practiced language (L1), we can derive analogous predictions about how sentence planning should differ when speakers use a less practiced language.

Incrementality in L2

Decades of research on second-language processing have shown that L1 and L2 production differ on several dimensions, e.g., production choices (word and structure choices), grammatical correctness, fluency, pronunciation, and processing speed (e.g., [Bosker et al., 2013](#); [Costa, 2005](#); [de Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2013](#); [Kroll & De Groot 2005](#); [Kroll & Gollan, 2014](#); [La Heij, 2005](#); [Piske et al., 2001](#); [Schoonbaert et al., 2007](#); [Schmidt, 1992](#); [Van Stutterheim, Flecken, & Carroll, 2013](#)). However, fundamental questions about the way speakers prepare utterances in a non-native language remain unanswered: while it is clear that the use of a second language may influence what speakers say and how quickly they are able to say it (even in advanced learners), it is unclear how it may also influence how they *prepare* to say it.

Our focus is thus not on L1-L2 differences in explicit linguistic knowledge or on the outcomes of production (correct utterances vs. errors), but on differences in the process of assembling sentences online. We exploit natural differences in speakers' ease of producing L1 and L2 sentences to assess how the level of *experience* (or “practice”) with a language shapes the time-course of formulation: speakers plausibly expect sentence production to be harder in L2 than in L1, and we hypothesize that this will result in differences in sentence planning. We test this hypothesis using a production task similar to previous L1 studies where speakers spontaneously described pictures of transitive events (like a dog chasing a mailman; [Griffin & Bock, 2000](#)).

A priori, we can expect at least two types of experience-driven planning strategies in L1 and L2. One possibility is that linguistic experience predicts the difficulty of linguistic encoding

alone (i.e., the ease of executing sentence-level processes) and does not influence planning strategies (i.e., the *coordination* of message-level and sentence-level processes). Specifically, slower linguistic encoding in L2 may result in a general slowing of production in L2: speakers may systematically use a linearly incremental or a hierarchically incremental planning strategy in both L1 and L2, and, regardless of the strategy they adopt, linguistic encoding should take longer in L2 than in L1. In this case, the prediction for an eye-tracked sentence production task would be that L1 and L2 formulation should not differ in early time windows (i.e., in windows where message-level and sentence-level encoding begin; 0-400 ms) but should differ in later windows (i.e., in windows where speakers engage primarily in sentence-level encoding). Here, L2 speakers should look longer at each character because lexical retrieval is slower in L2 than in L1.

A competing prediction is that experience with a language will have a stronger influence on formulation by shaping speakers' planning strategies in addition to predicting the difficulty of linguistic encoding. In this case, L1 and L2 formulation should differ from trial onset, reflecting differences in the coordination of message-level and sentence-level processes. Prior research supports two hypotheses, namely the possibility of shifts towards (a) a more linear strategy or (b) a more hierarchical strategy in L2.

One plausible way of planning sentences in a less practiced language is (a) to “break up” the planning process into smaller increments, consistent with *linearly incremental planning*. Support for this view comes from the finding that speakers use smaller planning windows when preparing sentences under cognitive load: smaller increments are easier to plan than larger increments, so the advantage of this strategy is that it allows speakers to “distribute” cognitive effort across an entire sentence (Ferreira & Swets, 2000; Wagner et al., 2010). Applying this strategy here, speakers may prefer to plan smaller increments (e.g., potentially one-character increments) in L2 in order to reduce processing load per increment. In an eye-tracked production task, linear planning should manifest as faster divergence of fixations to the two characters – i.e., a stronger preference for the subject character over the non-subject character – during early L2 than L1 formulation, indicating priority conceptual and linguistic encoding of this character alone (e.g., Gleitman et al., 2007). A disadvantage of this strategy is that speakers will need to continuously engage in planning of new conceptual and linguistic increments after speech onset. This may result in post-onset disfluencies and the need to make costly repairs if, for example, speakers change their mind about what they want to say or if they find a different way of doing so mid-sentence.

Alternatively, (b) speakers may opt for planning a larger message up front in L2 and thus may begin linguistic encoding only once they know, roughly, what they want to communicate (*hierarchically incremental planning*). In L1, shifts between linear and hierarchical planning occur from sentence to sentence when individual linguistic encoding processes become easier or harder to execute (Konopka & Meyer, 2014). The nature of these shifts suggests that speakers prefer to complete easier processes before harder processes. In L2, message-level encoding is arguably easier to complete than sentence-level encoding when it is independent of linguistic forms (i.e., when message-level encoding does not require linguistic mediation; see Footnote 1). So, when planning L2 sentences, speakers may prefer to encode a larger message with priority first (the “easy” process) and postpone linguistic encoding (the “harder” process). Thus, they may spend more time distributing their attention between two event characters immediately after picture onset in L2 than L1.

Hierarchical planning also circumvents the problem posed by Linear Incrementality of having to generate new increments of the message after speech onset. The fact that speakers must continue planning both conceptual and linguistic information after speech onset makes linear planning an “opportunistic” or “short-sighted” strategy. In the current example, it assumes that speakers may choose a particular character to be the starting point of their sentence (“dog” or “mailman”) based only on how easy this character is to encode. However, easy planning of the first increment does not imply easy planning of subsequent increments to complete the sentence: speakers may well run into difficulties after speech onset, resulting in mid-sentence pauses, disfluencies, and conceptual and linguistic repairs (e.g., *The mailman is running while...no, The dog is chasing the mailman*). Speakers can conduct repairs relatively quickly in L1, but beginning to encode an incomplete message linguistically is a risky strategy in L2 because speakers have fewer repair strategies available to them in their less practiced language (Sadri Mirdamadi & de Jong, 2015; Segalowitz, 2010). Thus, to reduce the likelihood of producing disfluent speech and making “last-minute” repairs mid-utterance in L2, speakers may prefer to do as much message-level planning up front as possible (0-400 ms, i.e., hierarchical planning) and leave primarily linguistic encoding to complete after 400 ms.

Current experiments

We test these accounts in four experiments using a well-established methodology to elicit and compare L1 and L2 descriptions of two-character transitive events (Konopka & Meyer, 2014) under different conditions. Spontaneous sentence production requires at least intermediate proficiency, so our participant sample was composed of university students with Dutch as their native language and high proficiency in English. Speakers in all experiments completed half of

the task in Dutch and half of the task in English, allowing perfect control over variables such as education level and motivation.

Experiment 1 established baseline differences in L1 and L2 formulation. Dutch and English are typologically very similar languages and use similar syntax in transitive sentences (same word order—SVO (active) and OVS (passive)—and no case marking; see [Norcliffe et al., 2014](#); [Sauppe et al., 2013](#)). Moreover, in both languages active structures are highly preferred over passives ([Bernolet, Hartsuiker, & Pickering, 2009](#); [Sadri Mirdamadi & de Jong, 2015](#); [Segaert, Wheeldon, & Hagoort, 2016](#)). Thus, we expect differences in formulation to arise primarily from differences in speakers' experience using the target languages rather than from differences in the grammatical properties of these languages. Next, Experiments 2 to 4 systematically manipulated the ease of completing different production processes to identify the sources of L1-L2 differences. Experiment 2 tested whether these differences were due to strategic shifts in the timing of *message-level* encoding by selectively facilitating encoding of the gist of target events in each language. Experiments 3 and 4 tested whether L1-L2 differences were due to differences in *sentence-level* encoding of two types of information—agent names (i.e., content words that convey non-relational information; Experiment 3) and verbs (i.e., content words that convey relational information; Experiment 4)—by selectively facilitating encoding of this information in each language. For a comprehensive assessment of L1-L2 differences, we compared the effects of these manipulations on (1) structure choice (i.e., *what* speakers said), (2) production speed (i.e., *how quickly* they said it), and (3) formulation (i.e., *how they prepared* to say it). Each experiment used a large number of target events (80 target pictures) to obtain stable estimates of planning strategies and experience-driven changes in formulation.

In our analyses of (1) structure choice and (2) production speed, we test the extent to which speakers produce similar outputs in L1 and L2. For a comparison against earlier studies, we assess the influence of two measures known to shape L1 production: *event codability* (or *action codability*) and *agent codability* ([Kuchinsky & Bock, 2010](#)). These measures reflect the ease of encoding and expressing the gist of the event (event codability) and naming the agent character (agent codability), and thus capture natural variation in the properties of every-day events. They also provide a fine-grained test of hierarchical and linear planning: effects of event codability on planning suggest sensitivity to *relational* properties of an event (a litmus test for hierarchical incrementality) and effects of agent codability suggest sensitivity to *non-relational* properties of an event (a litmus test for linear incrementality). Most importantly, in our analyses of (3) the time-course of formulation, we focus on L1-L2 differences in eye movements from trial onset until speech onset relative to the experimental manipulations.

In each experiment, we also obtained a basic assessment of L2 proficiency using a linguistic background questionnaire (based on [Marian et al.'s, 2007](#), LEAP-Q) and a vocabulary test (Lextale, available at <http://LexTale.com>; [Lemhöfer & Broersma, 2012](#)). We report LexTale scores to confirm a difference in participants' linguistic knowledge of Dutch and English. Vocabulary knowledge is a complex construct, but it remains a good measure of general proficiency even in advanced L2 users (e.g., [de Jong, et al., 2012](#); [Ellis, 2004](#); [Lemhofer & Broersma, 2012](#); [Luo, Luk, & Bialystok, 2010](#); [Zareva et al., 2005](#)).

Experiment 1

Experiment 1 assessed overall L1-L2 differences in sentence production and planning. Participants described pictured events spontaneously (e.g., a dog chasing a mailman), with no prior exposure to the pictures or to any L1 or L2 vocabulary. Thus, any differences observed between languages in structure choice, production speed and the time-course of formulation can be attributed to overall differences in speakers' experience using the two languages.

We expected no large L1-L2 differences in the content of target sentences, although lower proficiency in L2 may result in speakers producing lexically and structurally simpler sentences in L2 (see [de Jong et al., 2008](#)). Specifically, speakers may use a smaller range of lexical items and may prefer to use active syntax over passive syntax in English than in Dutch (even though syntax is shared across languages; [Hartsuiker, Pickering, & Veltkamp, 2004](#); also see [Kim & McDonough, 2008](#), and [Sadri Mirdamadi & de Jong, 2015](#)). In terms of production speed, we also expected participants to initiate speech more slowly in L2 than L1. Importantly, our main focus is on analyses of eye movements to test whether linguistic experience changes planning priorities. If formulation is more *linear* in L2 than in L1, then fixations to agents and patients in target pictures should diverge more quickly after picture onset in English than in Dutch sentences. In contrast, if formulation is more *hierarchical* in L2, speakers should be more likely to distribute their gaze between agents and patients for a longer period of time at the beginning of each trial in English than in Dutch sentences.

Method

Participants

Participants were 38 native Dutch speakers (23 female), with a mean age of 22 ($SD=3.2$, range=19-32), mostly university students at Radboud University Nijmegen, the Netherlands. Data from three additional participants were excluded due to technical problems resulting in the elimination of a large number of target trials, but no participant was excluded due to inability to complete the task in L2. Participants in all experiments were paid 8 Euros.

At the end of the experimental session, participants completed a questionnaire about their linguistic background and English-speaking experience as well as the LexTale test in L1 and L2. All participants used Dutch as their dominant language and spoke English as a non-native language. Students in the Netherlands are exposed to English and use English to communicate on a regular basis (see Table 1 for descriptive statistics for all experiments). As expected, LexTale scores were higher in Dutch than English (Table 2; $t[37]=6.11$), and were weakly correlated. Participants' self-reported English speaking-ability scores were strongly correlated with their English LexTale scores (Table 2), consistent with earlier studies (e.g., [Marian et al., 2007](#)). (Across all experiments, English LexTale scores were also correlated with the age at which participants reported beginning to learn English [r range: $-.16-.45$] and with the number of years that they reported speaking English [r range: $.34-.49$]).

Table 1. Participants' age of first acquisition of English, number of years using English, and self-reported use of English at the time of testing in Experiments 1-4.

Exp.	Age at the point of acquisition			Number of years using English			Current usage of English					
	<i>M</i>	<i>SD</i>	range	<i>M</i>	<i>SD</i>	range	At least...					total <i>n</i>
							several hours a day	once a day	once a week	once a month	once a year	
Exp.1	9.4	2.8	2-18	12.6	5	6-25	4	13	13	7	1	38
Exp.2	10.5	2.6	4-21	11	3.9	5-23	5	23	20	10	3	63
Exp.3	10.5	1.8	2-12	10.6	2.7	7-21	3	7	27	9	3	48
Exp.4	10.3	1.9	3-13	11.1	3.3	6-21	2	14	26	10	4	62

Table 2. Performance (accuracy) on the Dutch and English LexTale vocabulary test, English self-reported speaking-ability scores (on a 1-5 scale), and correlations between measures in each experiment.

Exp.	Dutch			English			English self-reports	Correlations	
	<i>M</i>	<i>SD</i>	range	<i>M</i>	<i>SD</i>	range	<i>M</i>	Dutch and Eng. LexTale	Eng. self-reports and Eng. LexTale
Exp.1	91%	6%	75-99%	80%	12%	60-96%	3.6	$r=.30, p=.07$	$r=.60, p<.0001$
Exp.2	92%	4%	79-99%	76%	12%	48-100%	3.2	$r=.40, p<.001$	$r=.40, p<.001$
Exp.3	91%	6%	70-99%	77%	12%	53-98%	3.3	$r=.20, ns$	$r=.46, p<.001$
Exp.4	91%	4%	78-99%	77%	11%	58-100%	3.1	$r=.55, p<.0001$	$r=.48, p<.0001$

Materials and design. The target items were 80 pictures of transitive events (Appendix A). There were also 160 filler items showing intransitive, transitive, or ditransitive events. Both target and filler items included events with human and non-human agents and patients. Fifty-four of the target events had human agents, so we expected a large majority of active sentences (see e.g., Branigan, Pickering, & Tanaka, 2008; Ferreira, 1994; Norcliffe et al., 2015). Because of participants' tendency to use basic-level referential terms, character names like *man/woman* or *boy/girl* were produced more often than more specific names (e.g., *barber* or *photographer*).

The experiment was divided into two counterbalanced blocks: half of the participants began the task in their L1 (Dutch) and then switched to L2 (English), and half began the task in their L2 (English) and then switched to L1 (Dutch). Target items in the first and second blocks were matched for Agent, Patient, and Event Codability based on ratings obtained in earlier studies using the same methodology (Konopka & Meyer, 2014; Konopka & Kuchinsky, 2015; Kuchinsky & Bock, 2010). Two mirror-reversed copies of each item were created to counterbalance the placement of the agent in the picture (left- vs. right-hand side of the picture). Counterbalancing this factor as well as block language and block order within participants and within items resulted in eight lists of stimuli. Experimental lists were then created by interspersing target events among fillers, with two filler items occurring between any two target items.

Procedure. Participants were instructed to produce one sentence to describe each pictured event, naming all event characters. Eye movements were recorded with an EyeLink 1000 tracker (1000 Hz sampling rate). Calibration was performed twice during the experiment, prior to each experimental block, using a 9-point calibration procedure. Each trial began with a fixation dot at the top of the screen. Participants pressed the space-bar while looking at this dot to display the next picture, and after they produced a description, the experimenter clicked with the mouse to advance to the next trial. At the end of the session, participants completed the LexTale task and the language questionnaire.

Sentence scoring. Event descriptions were scored as *actives* (SVO word order: e.g., "The dog chased the mailman"; "De hond zit de postbode achterna"), *passives* (OVS word order: "The mailman was chased by the dog", "De postbode wordt achterna gezeten door een hond"), *truncated passives* ("The mailman was chased", "De postbode wordt achterna gezeten"), or *other* responses (excluded from analysis). Sentence counts for all experiments are provided in Supplementary Materials A. Code switching was extremely rare, and sentences with code switches were excluded.

In English, speakers used either the present tense or the present progressive tense in their descriptions. In Dutch they consistently used the simple present tense; the equivalent of the progressive form (e.g., "De vrouw is de man aan het masseren") is uncommon (3% of sentences in each experiment). Most of the verbs speakers used in Dutch descriptions were simple and not separable (e.g., *slaan* [to hit]); less than 13% of active sentences in each experiment included morphologically complex and separable verbs (e.g., *achterna zitten* [to chase] where only the verb stem appears in second sentence position as for monomorphemic verbs: "De hond zit de postbode achterna") and English phrasal verbs (e.g., *to kick out*; "The waiter is kicking the chef out"). A small subset of the descriptions (less than 2%) included the verb "to try" or "to want to" (e.g., "De man probeert een kindje te redden")³.

Speech onsets were measured from picture onset to the onset of the first non-filler word of the sentence (i.e., excluding disfluencies like *um*, *uh*, etc.)⁴. Sentences with onsets over 5000 ms were excluded, as were any remaining sentences with onsets longer than three standard deviations above the overall mean onset time for that language. This left 1928 sentences for analysis (85% actives, 13% passives, 2% truncated passives, i.e. passives that omit the agent).

Codability Scoring. Codability scores (Shannon's entropy) were calculated for agents, patients, and actions (*Agent*, *Patient* and *Event* Codability respectively), independently for each language. Codability is a measure of how easy agents, patients, and actions are to describe, and is based on the distribution of participants' responses. Thus, agents, patients and events that are consistently named with a small set of words have low entropy and receive high codability scores. Conversely, the more variability there is in word choice (i.e., the higher the heterogeneity of speakers' responses), the lower the codability scores of that agent, patient or event are. Low codability scores indicate large between-participant differences in interpretation (conceptual codability) and word selection (linguistic codability).

Codability scores did not differ between Dutch and English for Events and Agents (both $t_s < 1$). This suggests little systematic difference in the range of lexical items that participants used in L1 and L2, and confirms that participants were highly proficient speakers of English. These measures were also strongly correlated across languages ($r=0.68$ for agents; $r=0.58$ for events),

³ The number of sentences with separable verbs was too small (less than 250 per experiment) to allow a separate analysis of eye movements. However, time-course analyses that did and did not include sentences with these verbs produced similar results in this experiment.

⁴ Speakers produced more sentence-initial disfluencies (i.e., disfluencies before the first character name) in English than in Dutch in all experiments: specifically, they produced early disfluencies in 12% and 23% of all Dutch and English sentences in Experiment 1, 10% and 18% Dutch and English sentences in Experiment 2, and 11% and 19% Dutch and English sentences in Experiments 3 and 4. Disfluency rates were lower and comparable across Dutch and English before the sentence verb and before the last character in the sentence (less than 8% of all sentences).

which warrants a direct comparison of performance across Dutch and English. Event and Agent Codability scores also did not differ between speakers with higher and lower Dutch and English LexTale scores.

Patient codability scores were strongly correlated in Dutch and English as well ($r=0.70$). However, effects of Patient codability on production were not assessed due to overwhelming production of active sentences and to a negligible contribution of properties of the patient to sentence planning in earlier studies (see [Ganushchak, Konopka, & Chen, 2016](#), for a direct test).

Analyses. In all experiments, we first conducted analyses of structure choice and speech onsets to test for differences in the L1 and L2 output. Analyses of structure choice were conducted using linear logit mixed effect models in R ([Jaeger, 2008](#); *Results section 1*). The models were run in two steps. The first analysis included Language and Block (categorical predictors) as well as LexTale scores (continuous predictor) for an initial assessment of how linguistic knowledge influenced production. Block was used instead of Trial Number for consistency with the analyses of eye movements (see below). The second analysis included Language, Block (categorical predictors) and Agent and Event Codability scores (continuous predictors; for clarity, LexTale scores were dropped these analyses). Categorical predictors were coded with Helmert contrasts; continuous predictors were converted to z-scores. Analyses of speech onsets were conducted using linear mixed models with the same variables for active sentences ([Baayen, Davidson, & Bates, 2008](#); *Results section 2*). The models included all three-way interactions between the factors of interest as well as random slopes for all factors included additively to allow convergence. When models did not converge, we removed slopes iteratively, starting with the random effects accounting for the smaller amount of variance (models with and without slopes produced similar results). All effects were considered to be significant at $p<.05$, and we report theoretically important results that are marginally significant ($p<.10$) for completeness.

Growth Curve Analyses (GCA; [Mirman, 2014](#)) were run to compare the time-course of formulation in L1 and L2 sentences with active syntax (i.e., the large majority of sentences in all experiments; *Results section 3*). The number of passive sentences produced in each experiment was too small to permit an analogous analysis (but see van de Velde et al., 2014, for an assessment of the formulation of passive sentences in L1).

GCA is a multilevel regression that fits orthogonal polynomials to describe linear and nonlinear changes in target fixations over time, and is thus preferred to approaches such as quasi-logistic regressions. All analyses included an intercept, a Linear term (capturing the overall slope of fixations), a Quadratic term (describing the sharpness of a single peak in the

distribution of fixations), and a Cubic term (describing the sharpness of two peaks), as well as interactions of these terms with the experimental variables. All models also included random additive slopes for the three time terms. The experimental variables were Language (Dutch was the baseline) and Block (Block 1 was the baseline). Random slopes for these variables are listed in Table captions.

GCAs were performed on agent-directed fixations by participants and by items (transformed to empirical logits) in two time windows: early (0-400 ms) and late (400-2000 ms, i.e., until speech onset in L2). Due to aggregation across items, Block was used as a factor instead of Trial Number. Time windows were selected based on previous studies' categorization of windows capturing primarily message-level and sentence-level encoding respectively (Ganushchak et al., 2014, 2016; Griffin & Bock (2000); Konopka & Kuchinsky, 2015; Konopka & Meyer, 2014; Norcliffe et al., 2015; Sauppe et al., 2013). All results are reported in tables, and we discuss theoretically important effects in the text (interactions between the experimental variables and the Linear, Quadratic, and Cubic time terms that are reliable by participants and by items).

To preview, the time-course of formulation was starkly different in the two blocks in all experiments, so we report separate analyses for each block first, followed by a joint analysis (where we comment only on interactions with the Block variable). While these differences between blocks were not predicted, they are complementary to the main predictions of the paper.

Results

(1) *Structure choice*

Effects of Language, Block, and Proficiency. The first model predicting structure choice included a three-way interaction between Language, Block, and LexTale scores (Dutch scores for Dutch sentences, English scores for English sentences), random by-participant slopes for Language and Block, and random by-item slopes for Language and LexTale scores. This model showed a weak main effect of Language ($\beta=0.86$, $z=1.90$, $p=.06$): speakers were more likely to produce sentences with active syntax – i.e., with “easy” structures – in English than in Dutch (*Figure 1a*). Structure choice was also modulated by LexTale scores ($\beta=0.75$, $z=2.50$, for the interaction of Language with LexTale): speakers with lower English scores were more likely to use active syntax and thus to avoid the dispreferred passive construction in English.

Effects of Codability. Effects of Event and Agent Codability were evaluated in a model that included a three-way interaction between Language, Block and Event Codability, and between Language, Block and Agent Codability (with random slopes for Language, Block, and Agent Codability; see Supplementary Material B for by-item plots). There were no interactions

with Language and only a weak interaction between Block and Event Codability ($\beta=.38$, $z=1.86$, $p=.06$): higher-codability events were more likely to be described with passive sentences in Block 2 than Block 1. There was also a weak effect of Agent Codability as speakers produced more passive sentences to describe events with harder-to-name agents in Dutch in Block 1; however, this effect was reduced by the addition of random slopes for Event Codability.

(2) *Speech onsets*

Speech onsets were shorter in active than passive sentences in both Dutch (1809 vs. 1948 ms) and English (1974 vs. 2147 ms), consistent with earlier studies ([Sadri Mirdamadi & de Jong, 2015](#)). Further analyses were run exclusively on active sentences.

Effects of Language, Block and Proficiency. The first model included a three-way interaction between Language, Block, and LexTale scores (with random by-participant slopes for Language, and random by-item slopes for all factors). Speakers were faster to produce active sentences in Dutch than in English ($\beta=173$, $t=3.87$; *Figure 2a*), and no other effects reached significance.

Effects of Codability. Next, speech onsets were examined in a model including all three-way interactions between Language, Block, Event Codability, and Agent Codability (with all random slopes). Speech onsets were faster for events with easy-to-name agents in both L1 and L2 ($\beta=97$, $t=3.80$), consistent with earlier demonstrations of character accessibility on speech onsets. The effect of Agent Codability was stronger in high-codability events than low-codability events ($\beta=-43$, $t=-2.11$, for the interaction of Event and Agent Codability), but there were no interactions with Language.

(3) *Time-course of formulation*

First Block

Immediately after picture onset (0-400 ms), speakers fixated agents and patients at similar rates, showing only a small preference for the agent in both L1 and L2 (*Figure 3a*). Time-course analyses carried out on this time window confirmed that there were no statistical differences between L1 and L2 (there were no interactions between Language and the Linear, Quadratic, or Cubic terms of the Time variable, except in the by-item analysis; Table 3a, *Figure 3a, 3c*).

Differences between languages emerged in the second time window (400-2000 ms). Speakers began fixating the agent preferentially right after 400 ms when using their L1, indicating that they began encoding this character linguistically at this point in time. In contrast, they continued distributing their gaze between the agent and the patient for another 200 ms when using their L2, and began fixating the agent preferentially only 600 ms after picture onset. This

delay in the rise of agent fixations in L2 shows that speakers dedicated more time to pre-linguistic encoding in L2, and is consistent with *Hierarchical Incrementality* in L2.

Once speakers shifted their attention to the agent character, they continued fixating the agent preferentially until approximately 1000 ms and then began moving their gaze away from this character earlier in L1 than in L2. Thus, the entire distribution of agent fixations in L2 was shifted to the right compared to L1. As this time window is argued to correspond to linguistic encoding, differences in the length of agent fixations indicate that lexical retrieval was initiated later and completed later in L2 than in L1.

These differences are captured by an interaction of Language with the Linear time term (Table 3b, *Figure 3c*): there were fewer agent-directed fixations at the beginning of this time window in English than in Dutch, and more fixations at the end of this time window in English than in Dutch. In the by-item model, there was also an interaction between Language and the Quadratic term, indicating a taller and sharper peak of agent fixations in English than in Dutch.

Table 3.

GCA results for active sentences in Experiment 1. All models included by-participant random slopes for Language. Statistically significant effects are highlighted in all tables.

Effect	By-participants			By-items		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
<i>First block</i>						
<i>a) 0-400 ms</i>						
Intercept	-1.38	.08	-18.00*	-1.00	.09	-11.25*
Linear	8.13	.33	24.75*	5.43	.34	16.05*
Quadratic	-2.49	.26	-9.64*	-.90	.20	-4.53*
Cubic	-1.17	.22	-5.32*	-1.32	.18	-7.51*
Language	.02	.15	.10	.01	.08	.10
Linear * Language	-.16	.66	-.24	-.03	.09	-.39
Quadratic * Language	.46	.52	.89	.25	.08	2.94*
Cubic * Language	-.41	.44	-.94	-.12	.08	-1.41
<i>b) 400-2000 ms</i>						
Intercept	.20	.06	3.48*	.22	.06	3.53*
Linear	-2.34	.46	-5.05*	-1.19	.44	-2.70*
Quadratic	-3.91	.26	-15.32*	-3.14	.47	-6.72*
Cubic	2.10	.35	5.99*	1.42	.40	3.56*
Language	.08	.11	.72	.13	.05	2.59*
Linear * Language	2.53	.93	2.72*	2.26	.10	23.02*
Quadratic * Language	-.60	.51	-1.18	-.57	.10	-5.83*
Cubic * Language	-.03	.70	-.05	-.01	.10	-.10

*Second block**c) 0-400 ms*

Intercept	-1.16	.08	-14.34*	-.82	.10	-8.39*
Linear	8.34	.42	19.89*	5.48	.34	15.96*
Quadratic	-2.52	.24	-10.72*	-.85	.23	-3.74*
Cubic	-.86	.24	-3.63*	-1.37	.17	-8.12*
Language	.09	.16	.55	-.03	.07	-.44
Linear * Language	.24	.84	.29	-.06	.10	.64
Quadratic * Language	-.68	.47	-1.45	-.29	.09	-3.12*
Cubic * Language	.69	.48	1.46	.42	.09	4.60*

d) 400-2000 ms

Intercept	.36	.04	9.23*	.29	.05	5.34*
Linear	-1.66	.41	-4.08*	-.92	.52	-1.78
Quadratic	-3.38	.40	-8.44*	-2.92	.47	-6.21*
Cubic	.43	.34	1.26	.26	.31	.85
Language	.09	.08	1.16	.13	.05	2.57*
Linear * Language	-.23	.81	-.29	-.37	.10	-3.67*
Quadratic * Language	.80	.80	1.00	1.14	.10	11.31*
Cubic * Language	.31	.68	.45	.28	.10	2.84*

* $p < .05$ † $p < .10$

Second Block

Speakers showed a stronger preference for fixating a single character immediately after picture onset in Block 2 regardless of Language: they directed their gaze to the agent very quickly before 400 ms in both L1 and L2 (*Figure 3b, 3d*). This pattern indicates a global shift towards *Linear Incrementality*, as it shows priority encoding of one character alone early in the formulation process. There were no interactions with Language in the by-participant analysis (Table 3c).

Between 400 and 2000 ms, L1-L2 differences were numerically smaller than in Block 1, and all effects reached significance only in the by-item analysis (Table 3d). In general, there were more agent-directed fixations at the beginning of this window in L2 than L1, and even though speakers began encoding the L2 agent with priority much earlier than in Block 1, they still looked longer at the agent when producing L2 than L1 sentences.

Joint analysis of the First and Second Block

To confirm the presence of a within-experiment shift in formulation patterns, a joint analysis was conducted to compare the distributions of agent-directed fixations across blocks (see Supplementary Material C for the output).

Between 0 and 400 ms, fixations on the agent were more sustained (i.e., less variable) in Block 2 than Block 1 (interaction between Block and the Cubic term), confirming a reliable shift

towards priority encoding of the agent over the course of the experiment. In the 400-2000 ms time window, agent fixations dropped more slowly (interaction with the Linear term), had a more shallow peak (interaction with the Quadratic term) and were more sustained in Block 2 than Block 1 (interaction with the Cubic term). Thus, speakers spent more time looking at the agent and began adding the patient character to the developing sentence later in Block 2 than in Block 1. There was only one interaction with Language (Language, Block and the Linear term in the by-item analysis), as the fastest drop in agent fixations was observed in Block 1 in Dutch sentences.

Discussion

In Experiment 1, the outcome of L2 production (i.e., structure choice and sentence onsets) and the time-course of L2 formulation (i.e., eye movements before speech onset) were sensitive to different variables than in L1.

Analyses of structure choice showed production patterns typical of L1 when speakers were using Dutch (Konopka & Meyer, 2014) and important differences when speakers were using English. First, speakers were more likely to use both active and passive syntax in L1, but preferred to use active syntax in L2 (*Results section 1*): this is consistent with the expectation that speakers should show greater structural flexibility in their native language but should select “easier” structures in their less practiced language. Further, passive syntax was used more frequently in speakers with higher English vocabulary scores.

The target events varied in Event and Agent Codability. Speakers generally prefer to begin sentences with accessible subjects, but Agent Codability predicted structure choice very weakly in L1 and not in L2. Poorer sensitivity to codability measures in L2 suggests that selection of active syntax in L2 may have occurred, to a large extent, “by default” and independently of the higher-level conceptual and linguistic factors that influence L1 production.

Second, speakers were slower to begin sentences in their L2 than in their L1 (*Results section 2*). As expected, there were strong effects of Agent codability in L1 and L2: Dutch sentences with easy-to-name agents were initiated more quickly than sentences with harder-to-name agents.

Finally and most importantly, time-course analyses showed a different pattern of L1 and L2 formulation (*Results section 3*). In Block 1, formulation in both languages began with fixations being distributed relatively equally between the event characters in the first 400 ms of picture viewing, suggesting that speakers generally preferred to devote time and resources to encoding a larger preverbal message before starting linguistic encoding in both L1 and L2. However, after 400 ms, speakers began fixating the agent character preferentially when

preparing L1 sentences, but did so only after 600 ms when preparing L2 sentences. In other words, speakers delayed the onset of linguistic encoding and thus were more likely to temporally separate message-level encoding from sentence-level encoding in L2 than in L1. The direction of this effect is consistent with Hierarchical Incrementality.

We believe that alternative explanations for the delayed rise of agent-directed fixations in L2 are unlikely. For example, one alternative explanation is that speakers might have delayed the onset of linguistic encoding of the agent because they were searching for a different starting point for their sentences (this pattern is, in fact, observed in L1 active sentences when agents are hard to name: speakers quickly shift their attention to the patient to determine if this character can serve as an alternative sentence subject; [Konopka & Meyer, 2014](#)). This is unlikely in the current experiment because speakers generally preferred to avoid passive syntax and did not show sensitivity to properties of the agent when choosing sentence structure in L2. It is also unlikely that they preferred to start encoding the name of the patient early on because the length of the L1-L2 delay in early fixations was much shorter (approximately 200 ms) than is needed to retrieve a character name and because speakers generally start encoding the patient name much later in the formulation process (see e.g., [Ganushchak et al., 2016](#)).

Our explanation of L1-L2 gaze differences in Block 1 being due to strategic differences in L1 and L2 formulation also receives support from another finding in the same experiment: the within-participant shift towards linearly incremental planning from Block 1 to Block 2 in both languages. This shift was observed irrespective of the language used in Block 1. In fact, due to baseline differences in Block 1 (more agent-directed fixations in L1 than L2), there were more agent-directed fixations in the 0-400 ms time window of Block 2 in L2 than in L1: participants who started the experiment speaking in their L1 and then switched to L2 showed stronger linear planning in Block 2 than participants who started the experiment speaking in their L2 and then switched to L1. Importantly, this result shows that speakers *are* able to plan L2 sentences in a linearly incremental manner and that their choice to do so can be modulated by experience (in this case, experience with the task).

In the next three experiments, we tested the validity of these conclusions by comparing L1 and L2 formulation under different conditions. To zoom in on individual production processes and identify the processes responsible for L1-L2 differences, Experiments 2-4 compared L1-L2 formulation after manipulating the ease of completing *message-level* encoding (Experiment 2), and *sentence-level* encoding of the *agent* name (Experiment 3) and of the *verb* (Experiment 4).

Experiment 2

In Experiment 1, we hypothesized that the early L1-L2 difference in the time-course of formulation was due to speakers taking more time to encode the preverbal message before beginning linguistic encoding in L2. If this explanation is correct, facilitating message-level encoding should reduce the early L1-L2 difference. We test this hypothesis in Experiment 2 by increasing speakers' familiarity with the target events prior to the main experiment.

Speakers described the same events as in Experiment 1, but before the start of each block, they received a preview of half of the target pictures in that block. The targets were intermixed with filler pictures and were presented at a fast pace (500 ms per picture). This presentation time is more than sufficient for speakers to encode the gist of an event and to identify event characters (presentation times as short as 100 ms allow for coarse event identification; [Dobel, Gumnior, Bolte, & Zwitserlood, 2007](#); [Glanemann, Zwitserlood, Bolte, & Dobel, 2016](#); [Hafri, Papafragou, & Trueswell, 2013](#); [Henderson & Ferreira, 2004](#)) and to begin to activate lexical content (e.g., [Chabal & Marian, 2015](#)), but it is not sufficient to *complete* linguistic encoding of content words (character names or the sentence verb). Fast presentation times also reduce the possibility of speakers continuing to process a picture after its offset because each new stimulus acts as a mask.

The Preview manipulation was not expected to have a strong influence on structure choice, although familiarity with a message may allow participants to produce more complex descriptions (e.g., descriptions with a wider range of lexical items or more difficult structures). Instead, we expected effects of Preview on the time-course of formulation: if speakers prefer to delay linguistic encoding in L2 to devote more time and resources to early message-level encoding (0-600 ms), there should be a smaller L1-L2 difference in early eye movements in the Preview than the No Preview condition. If the reason for the delay in linguistic encoding in L2 lies elsewhere, the Preview manipulation should not change the time-course of formulation.

Method

Participants

Participants were 63 native Dutch speakers (44 female), with a mean age of 21 ($SD=2.44$, range=18-28), from the same participant pool as Experiment 1. Data from three additional participants were excluded due to technical problems.

Table 1 lists descriptive statistics for this sample. LexTale scores were again higher in Dutch than in English (Table 2, $t[62]=11.18$), and were strongly correlated. Participants' self-reported English speaking-ability scores were also strongly correlated with English LexTale scores (Table 2).

Materials, design, and procedure. The materials were the same as in Experiment 1. The experiment was again divided into two counterbalanced blocks, one to be completed in L1 and

one in L2. Half of the items (i.e., 20 pictures) from each block were selected to be previewed (counterbalanced across lists). The previewed and non-previewed events from each block had similar Event and Agent Codability scores, based on earlier experiments conducted with these materials and with speakers from the same participant pool. Counterbalancing the previewed pictures, their left-right orientation, block language, and block order within participants and within items resulted in 16 experimental lists.

The preview was provided immediately before each block: participants saw 20 target pictures and 22 randomly chosen fillers from that block. They were told they would see each picture for 500 ms, and that seeing these pictures would facilitate performance in the main task.

Sentence scoring, codability scoring, and analyses. Applying the same scoring criteria as in Experiment 1 left 2359 sentences for analysis (84% actives, 12% passives, 4% truncated passives).

Event, Agent, and Patient Codability scores were computed for sentences in the Preview and No Preview conditions separately. All codability scores were again highly correlated in Dutch and English ($r=0.40$ for events in the Preview and No Preview conditions; $r=0.32$ and $.53$ for agents in the two conditions; $r=0.51$ and $.30$ for patients in the two conditions; all $ps<.05$). Codability scores did not differ in the two conditions in Dutch. In English, Event Codability scores were lower in the Preview condition (i.e., speakers used a wider range of verbs) than in the No Preview condition, and lower than Dutch Event Codability in the Preview condition. Agent Codability ratings were lower in Dutch than in English (i.e., speakers used a wider range of agent names in Dutch). Event Codability scores did not differ between speakers with higher and lower Dutch and English LexTale scores. Agent Codability scores, however, showed more variability in speakers with higher Dutch and English scores.

Analyses were carried out as in Experiment 1, with the addition of the Preview variable (No Preview was the baseline in GCAs).

Results

(1) Structure choice

Effects of Language, Block, Preview, and Proficiency. The first model included the three-way interaction between Block, Language, Preview and LexTale scores, but showed no main effects or interactions between these variables (all $zs<1.2$; *Figure 1b*).

Effects of Codability. The second model evaluated the effects of Codability by testing all three-way interactions between Language, Preview, Event Codability, and Agent Codability (with random slopes for Language and Block; Block was included as an additive factor).

There was a weak three-way interaction between Language, Preview, and Event Codability ($\beta=.73$, $z=1.90$, $p=.057$). When describing high-codability events in Dutch, participants were more likely to produce passive sentences in the Preview condition, suggesting that increasing familiarity with the events also increased the likelihood of speakers describing those events with more complex structures. Preview had no effect on structure choice for low-codability events in Dutch and no systematic effect on structure choice in English⁵.

(2) *Speech onsets*

Speech onsets were shorter in active than passive sentences in both Dutch (1833 vs. 2019 ms) and English (2012 vs. 2220 ms). Further analyses were run on active sentences.

Effects of Language, Preview, Block, and Proficiency. The first model tested for the three-way interactions between Preview, Language, Block, and LexTale scores (with all random slopes), and showed that participants were faster to start speaking in Block 1 than Block 2 ($\beta=123$, $t=3.41$), and faster in Dutch than in English ($\beta=187$, $t=5.32$; *Figure 2b*). There were no effects of Preview (suggesting that speakers did not systematically begin retrieving lexical material during the picture preview) and no effect of LexTale scores.

Effects of Codability. Next, effects of Codability were examined in a model including all three-way interactions between Language, Preview, Event Codability, and Agent Codability (with all random slopes). Block was included as an additive factor. This model showed again that onsets were faster in Dutch than English and faster in Block 1 than Block 2. Importantly, onsets were again shorter in sentences beginning with easy-to-name agents in both L1 and L2, and there was an interaction between Preview, Event Codability and Agent Codability ($\beta=70$, $t=2.75$). Speech onsets were generally faster in the Preview than the No Preview condition, except for high-codability events with hard-to-name agents, suggesting that speakers engaged in more extensive early planning for these descriptions.

(3) *Time-course of formulation*

First Block

Before 400 ms, fixations to agents rose more slowly in the Preview than the No Preview condition (interaction of Preview with the Linear term; Table 4a; *Figure 4a, 4c*), so speakers were more likely to distribute their attention between agents and patients in the Preview

⁵ The model also showed a three-way interaction between Language, Event Codability, and Agent Codability ($\beta=-.50$, $z=-1.85$, $p=.066$). In Dutch, speakers produced more active sentences when describing high-codability events, both when agents were easy and hard to name. In low-codability events, there were more active sentences produced to describe events with easy-to-name agents and more passive sentences produced to describe events with hard-to-name agents (Konopka & Meyer, 2014; Kuchinsky & Bock, 2010). There were no codability effects in English. Thus, properties of the events and agents influenced structure choice only in the speakers' L1.

condition in both Dutch and English. This was likely due to the fact that speakers were verifying that a given picture had indeed been previewed or because they wanted to *complete* message-level encoding of the entire event before starting linguistic encoding; both possibilities result in speakers dedicating more time and processing resources to encoding the entire event rather than a single character in the Preview than the No Preview condition.

In the No Preview condition, there were more agent fixations in Dutch than English, replicating Experiment 1, but this effect reached significance only in the by-item analysis (interaction between Language and Preview). Agent fixations also had a steeper slope in Dutch than English, but this effect was again significant only in the by-item analysis (interaction between Language and the Linear term).

Table 4.

GCA results for active sentences in Experiment 2. All models included by-participant or by-item random slopes for Language and Preview, except for models (c) and (d) which only included by-participant random slopes for Preview.

Effect	By-participants			By-items		
	Est.	SE	t	Est.	SE	T
<i>First block</i>						
<i>a) 0-400 ms</i>						
Intercept	-.85	.04	-19.17*	-.66	.07	-9.71*
Linear	4.68	.20	22.96*	3.59	.21	17.34*
Quadratic	-2.19	.17	-12.74*	-1.51	.18	-8.49*
Cubic	.08	.13	.64	-.13	.10	-1.29
Language	-.04	.09	-.50	-.01	.07	-.22
Preview	-.09	.02	-1.02	-.14	.06	-2.41*
Language * Preview	.03	.18	.19	.08	.03	3.31*
Linear * Language	-.34	.41	-.82	-.39	.08	-5.02*
Quadratic * Language	.41	.34	1.19	.11	.08	1.43
Cubic * Language	-.14	.26	-.53	-.08	.08	-1.01
Linear * Preview	-.30	.09	-3.14*	-.18	.09	-2.25*
Quadratic * Preview	.02	.09	.18	-.13	.08	-1.65
Cubic * Preview	-.07	.08	-.88	-.07	.08	-.90
Linear * Lang * Preview	-.06	.19	-.34	-.12	.16	-.76
Quadratic * Lang * Preview	.00	.18	.01	-.17	.15	-1.09
Cubic * Lang * Preview	-.11	.17	-.66	.07	.15	.46
<i>b) 400-2000 ms</i>						
Intercept	.19	.04	4.94*	.15	.04	3.98*
Linear	-1.89	.40	-4.77*	-1.29	.38	-3.39*
Quadratic	-3.55	.36	-9.92*	-2.93	.34	-8.50*
Cubic	.68	.23	2.99*	.54	.31	1.73

Language	.02	.08	.28	-.02	.04	.55
Preview	-.05	.05	-1.01	.00	.05	.02
Language * Preview	-.14	.10	-1.38	-.09	.01	-6.11*
Linear * Language	-.66	.80	-.83	-.54	.09	-6.24*
Quadratic * Language	-.04	.72	-.06	-.27	.09	-3.10*
Cubic * Language	.62	.46	1.36	.77	.09	8.92*
Linear * Preview	-.32	.09	-3.75*	-.18	.09	-2.12*
Quadratic * Preview	-.24	.09	-2.83*	-.29	.09	-3.38*
Cubic * Preview	1.13	.09	13.12*	1.03	.09	12.06*
Linear * Lang * Preview	-.39	.17	-2.30*	-.68	.17	-3.99*
Quadratic * Lang * Preview	.13	.17	.77	.11	.17	-.06
Cubic * Lang * Preview	.16	.17	.94	.48	.17	2.78*

*Second block**c) 0-400 ms*

Intercept	-.82	.06	-12.72*	-.65	.07	-9.33*
Linear	4.73	.24	19.76*	3.70	.21	17.94*
Quadratic	-1.83	.16	-11.70*	-1.23	.18	-6.71*
Cubic	.12	.14	.85	-.12	.09	-1.32
Language	.15	.13	1.15	.09	.07	1.39
Preview	.04	.10	.44	.10	.06	1.50
Language * Preview	-.11	.20	-.54	-.07	.02	-3.08*
Linear * Language	.80	.48	1.66	.61	.07	8.14*
Quadratic * Language	-.14	.31	-.46	.04	.07	.54
Cubic * Language	-.17	.28	-.61	-.17	.07	-2.39*
Linear * Preview	-.08	.09	-.89	.04	.08	-.56
Quadratic * Preview	-.12	.09	-1.45	-.06	.07	-.82
Cubic * Preview	.08	.08	1.06	.09	.07	1.30
Linear * Lang * Preview	-.45	.18	-2.51*	-.33	.15	-2.23*
Quadratic * Lang * Preview	.17	.17	1.00	.21	.15	1.43
Cubic * Lang * Preview	.13	.16	.83	.44	.14	3.01*

d) 400-2000 ms

Intercept	.30	.03	8.72*	.26	.05	5.31*
Linear	-1.46	.44	-3.29*	-1.23	.35	-3.49*
Quadratic	-2.80	.25	-11.06*	-2.42	.34	-6.91*
Cubic	.80	.26	3.08*	.69	.26	2.67*
Language	.10	.07	1.48	.07	.04	1.85†
Preview	.06	.04	1.43	.02	.04	.43
Language * Preview	.12	.09	1.33	.02	.01	1.32
Linear * Language	1.08	.88	1.22	1.07	.08	12.62*
Quadratic * Language	.87	.51	1.71†	.56	.08	6.56*
Cubic * Language	-.33	.52	-.63	-.56	.08	-6.57*
Linear * Preview	.01	.09	.05	.14	.08	1.66
Quadratic * Preview	-.17	.09	-1.89†	.20	.09	2.33*
Cubic * Preview	.00	.09	.05	-.05	.09	-.54
Linear * Lang * Preview	.19	.18	1.04	-.15	.17	-.91
Quadratic * Lang * Preview	.01	.18	.08	-.18	.17	-1.03
Cubic * Lang * Preview	.20	.18	1.11	.18	.17	1.04

* $p < .05$ † $p < .10$

After 400 ms, speakers began directing their gaze to the agent. Following from the previous time window, there were fewer agent fixations in the Preview than the No Preview condition at 400 ms, but then agent fixations rose and fell more quickly in the Preview than the No Preview condition: agent fixations had a more negative slope (interaction of Preview and the Linear term), a sharper peak (interaction of Preview with the Quadratic term), and were less sustained (interaction of Preview with the Cubic term) in the Preview than the No Preview condition (Table 4b; *Figure 4a, 4c*). It is unlikely that encoding of the agent took less time in the Preview condition; instead, this effect shows that speakers were prepared to begin adding the patient to the developing sentence earlier when describing a familiar event. This pattern of “looking ahead” to the next character suggests that sentence planning was more likely to be guided by a larger message-level framework in the Preview than the No Preview condition.

The only interaction with Language that was reliable by participants and by items included the Preview variable and the Linear term. There were two L1-L2 differences driving this interaction. First, at the beginning of this time window (400-2000 ms), speakers looked more often at the agent in the No Preview condition in English than in Dutch. This is different from Experiment 1, where speakers took longer to direct their gaze to the agent in English. By comparison, agent-directed fixations in the Preview condition were nearly identical in Dutch and English, as predicted. This confirms that there is little difference in the onset of linguistic encoding in L1 and L2 when speakers describe familiar events.

Second, by the end of this time window (2000 ms), there was a sharper drop in fixations in the Preview than No Preview condition in English: participants began shifting their gaze away from the agent and towards the patient earlier when speaking English than when speaking Dutch. Again, it is unlikely that encoding of the agent took less time in L2 than in L1; instead, speakers were more likely to “look ahead” when preparing L2 sentences.

Second Block

As in Experiment 1, there was a shift towards linearly incremental planning in Block 2: speakers directed their attention to agents more quickly after picture onset (0-400 ms) than they did in Block 1, and this shift was stronger in the No Preview condition in English (i.e., after speakers completed Block 1 in Dutch) than in Dutch (interaction of Language, Preview and the Linear term; Table 4c; *Figure 4b, 4d*).

After 400 ms, agent fixations had a more shallow peak in English than Dutch (interaction of Language with the Quadratic term; Table 4d; *Figure 4b, 4d*). This is similar to Experiment 1 and shows that speakers fixated the agent for longer in L2. The by-item analysis also showed a more shallow downward slope in English than Dutch fixations over this window (interaction

between Language and the Linear term). Effects of Preview were inconsistent across by-participant and by-item analyses.

Joint analysis of the First and Second Block

The output of the joint analysis is presented in Supplementary Material D. The 0-400 ms analysis showed that the distribution of agent fixations was more peaked in Block 1 than Block 2: speakers briefly looked away from the agent in Block 1 but continued fixating the agent in Block 2, consistent with a shift towards Linear Incrementality (interactions of Block with the Quadratic term). The analysis also showed a difference in slopes of agent fixations across blocks in Dutch and English: there were more agent fixations in Block 2 in English than in Dutch (interaction of Block, Language, and the Linear term). Finally, there was a difference in slopes in agent fixations across blocks in the Preview and No Preview conditions: speakers were less likely to fixate the agent in the Preview condition in Block 1, but more likely to fixate the agent in the Preview condition in Block 2 (interaction between Block and Preview, by-participant interactions of Block, Preview, and the Linear term). This confirms the presence of a shift towards linearly incremental planning in Block 2.

The 400-2000 ms analysis showed two effects of interest with Block. First, there was a more shallow peak to the distribution of agent fixations in Block 2 than Block 1 (interaction of Block with the Quadratic term), as speakers were more likely to “look ahead” to the patient in Block 1 and less likely to do so in Block 2. Second, speakers were also less likely to “look ahead” in the Preview condition in Block 2: agent fixations had a more shallow slope in the Preview condition in Block 2 than Block 1 (interaction of Block, Preview, and the Linear terms) and were more sustained in the Preview condition in Block 2 than Block 1 (interaction of Block, Preview, and the Cubic term). These effects are consistent with linearly incremental planning, as they show that speakers engaged in little advance planning beyond the first character in Block 2.

Discussion

In Experiment 2, speakers produced more sentences with “easy” structures (actives) in English than Dutch again and structure choice was sensitive to Preview only in Dutch high-codability events (*Results section 1*). These findings are consistent with Experiment 1 and show systematic effects of event properties on structure choice primarily in the speakers’ native language. Sentences beginning with easy-to-name agents had faster onsets in L1 and L2 again, but the picture preview had a small effect on speech onsets (*Results section 2*).

More importantly, time-course analyses showed strong effects of Preview on sentence formulation (*Results section 3*). In the No Preview condition of Block 1, there were again fewer agent-directed fixations in English than in Dutch (as in Experiment 1), although this difference

did not reach significance. In the Preview condition, speakers were more likely to distribute their attention between agents and patients in the first time window (0-400 ms) when preparing both Dutch and English sentences. This pattern is generally consistent with Hierarchical Incrementality as it indicates more extensive early encoding of the entire event, and thus formulation of a larger message-level framework in the Preview than the No Preview Condition.

There was also a clear effect of Preview in the second time window of Block 1 (400-2000 ms): in the Preview condition, speakers directed their attention to the agent more quickly after 400 ms and then began looking away from this character more quickly after 1000 ms. On [Griffin and Bock's \(2000\)](#) account, speakers encode a message framework that includes information about two characters in the early time window (0-400 ms) and this framework directs shifts of gaze to the two characters (the subject character and then the object character) in the order of mention for the purposes of linguistic encoding after 400 ms. Thus, by analogy, we propose that in the Preview condition of the current experiment, speakers likely showed earlier, more efficient gaze shifts to the agent and then the patient because these shifts were driven by the message framework encoded shortly after picture onset.

The effects of Language in the second time window were more complex. As expected, the distributions of agent-directed fixations in Dutch and English were very similar in the Preview condition immediately after 400 ms. Thus, when L2 speakers are familiar with an event, they no longer show a strategic delay in the onset of L2 linguistic encoding. Further, agent-directed fixations dropped more quickly in English than in Dutch towards the end of this window in the Preview than the No Preview condition. Lexical retrieval takes longer in L2 than in L1 (and it is unlikely that the Preview manipulation selectively facilitated encoding of the agent character), so this effect does not reflect faster encoding of the agent character in L2. Instead, it is more likely that this effect is also a consequence of L2 formulation being driven by a message framework. The fact that speakers shifted their attention towards the second character earlier in English than in Dutch suggests that speakers benefited more from the same Preview manipulation in L2 than in L1. This may be because speakers were already more likely to engage in extensive message planning in L2 and the Preview manipulation merely enhanced existing L1-L2 differences, or because the picture preview facilitated the completion of any conceptual processing for the patient still to be carried out at this stage. Both explanations are consistent with Hierarchical Incrementality.

The No Preview condition showed an unexpected pattern: speakers directed their attention to the agent more quickly in English than in Dutch. This pattern was observed once again in Experiment 4 (see below) where previewing verbs also resulted in participants initiating

linguistic encoding more quickly after 400 ms in English than in Dutch. These results thus suggest that the information that speakers extracted from the fast picture preview before each Block was likely information needed to encode the relationship between the two event characters. We return to this point in the General Discussion.

Finally, the results from Block 2 of this experiment showed a similar shift to Linear Incrementality as in Experiment 1, providing further support for the hypothesis that planning strategies change with experience with the task. In the first time window (0-400 ms) in the Preview condition of Block 2, speakers fixated the agent character very quickly; in the No Preview condition, speakers were more likely to adopt a linearly incremental planning strategy in English (i.e., after completing Block 1 in Dutch) than in Dutch (i.e., after completing Block 1 in English). In the second time window (400-2000 ms), interactions with Language reached significance only in the by-item analysis, indicating a high amount of variability in the data. Nevertheless, the results show that speakers fixated the agent for a longer period of time when preparing English sentences in Block 2 than Block 1. This is consistent with the expectation of longer lexical retrieval times in L2. It also provides further support for linearly incremental planning as it shows that the length of fixations depended primarily on the ease of lexical encoding rather than being guided by a message framework.

Experiment 3

Experiments 3-4 took a complementary approach to assessing the nature of early L1-L2 differences in formulation: instead of manipulating the ease of message-level encoding (as in Experiment 2), these experiments selectively facilitated the ease of *linguistic* encoding. If speakers prefer to postpone linguistic encoding in L2, which processes are responsible for this choice and present the largest processing bottleneck? If the early L1-L2 differences in eye movements observed in Experiments 1-2 were due to speakers delaying retrieval of L2 content words, then facilitating lexical retrieval should also reduce or eliminate this difference. Crucially, the two experiments together assess what type of information speakers encode with priority at the outset of formulation: Experiment 3 manipulated the ease of encoding the sentence-initial noun (i.e., the agent character name), and Experiment 4 manipulated the ease of encoding the sentence-medial verb.

Nouns and verbs map onto an important distinction in terms of message-level and sentence-level information and their implications for formulation (Konopka & Meyer, 2014). Character names convey *non-relational* information, i.e., they communicate the identities of event characters (e.g., *dog*, *mailman*) that are largely independent of their roles in the event. On the other hand, verbs convey *relational* information, i.e., they express the relationship between

event characters (e.g., *chasing*) and thus, by default, they require encoding of some information about the event roles of both characters (the chaser and the chasee). The degree to which speakers show sensitivity to relational and non-relational information in L1 and L2 can distinguish linearly incremental planning from hierarchically incremental planning.

Experiment 3 used a similar procedure to Experiment 2: speakers received a preview of character names, including names for half of the *agents* in the target events, before beginning the main experiment. Earlier studies also provided speakers with words they were to use during the experiment (e.g., Cai et al., 2011; Hartsuiker et al., 2004; Hwang & Kaiser, 2014); however, participants either received these words immediately prior to a target trial or they were trained to use specific words to describe specific pictures. Both methods reduce the complexity of the task by eliminating the need to search for suitable lexical items and may cue speakers to attend to specific areas in an event. Here, we use an implicit priming manipulation: we provided a noun preview at the beginning of the study and speakers were then free to use any words they liked in the main task. Facilitating agent naming was expected to result in more frequent use of active syntax in the Preview condition (replicating earlier accessibility effects on sentence structure) and in a shift towards linearly incremental planning (Konopka & Meyer, 2014). Importantly, if the processing bottleneck in L2 planning is in retrieving the first content word of the sentence, then facilitating retrieval of that word should result in a similar shift in planning strategies in L1 and L2. If the reason for the delay in linguistic encoding in L2 lies elsewhere, the Preview manipulation should not change the time-course of L2 formulation.

Method

Participants

Participants were 48 native Dutch speakers (31 female), with a mean age of 21 years ($SD=2.08$, range=18-26), from the same participant pool as in earlier experiments. Data from 11 additional participants were excluded due to technical problems⁶.

Table 1 lists descriptive statistics for this sample. LexTale scores were again higher in Dutch than English (Table 2, $t[46]=7.23$), but were not correlated. However, participants' self-reported English speaking-ability scores were strongly correlated with their English LexTale scores (Table 2).

Materials, design, and procedure. The materials were the same as in Experiment 1. The experiment was again divided into two counterbalanced blocks, one to be completed in L1 and

⁶ These problems were due to our initial temporary use of a new eye-tracker in this experiment. To avoid such problems, we switched back to our older eye-tracker.

one in L2. Half of the items from each block were selected to have previewed agent names, and the two halves had similar Event and Agent Codability scores. The selected agent names were the most frequency nouns used to describe agents in previous experiments. Counterbalancing the previewed agent names, the left-right orientation of the target pictures, block language, and block order within participants and within items resulted in 16 experimental lists.

The Preview before each block consisted of agent names for 20 of the target events and 22 randomly chosen nouns. Presentation rates were again quite fast (words appeared on the screen for 500 ms each), and participants were asked to read the words to themselves silently. They were told that familiarity with these words would facilitate performance in the main task.

Sentence scoring, codability scoring, and analyses. Applying the same scoring criteria as in previous experiments left 1784 sentences for analysis (87% actives, 11% passives, 2% truncated passives). Analyses were carried out as in previous experiments (*Results sections 1, 2, 3*).

Since the previewed nouns were the most common, but not the only possible agent names, the Preview manipulation could not ensure that speakers would use exclusively those nouns in their descriptions. In fact, speakers used the previewed nouns on approximately half of all previewed trials (e.g., speakers sometimes referred to an object falling onto a ship with the word “rocket” instead of using the previewed noun “bomb”), but also used the previewed nouns on trials designed to have no preview. Thus, analyses were carried out after recoding all responses into sentences that did and did not use the previewed nouns, irrespective of condition (i.e., a response was coded as “previewed” or “not previewed” based on speakers’ word choices): there were 877 sentences with previewed agent names and 865 sentences with non-previewed agent names. Speakers were as likely to use the previewed agent names in Dutch (50% of 936 sentences) as they were in English (51% of 806 sentences).

Agent Codability was calculated across the new recoded conditions, and so was naturally higher in the Preview condition (which had perfect name agreement) than the No Preview condition (where speakers used a range of non-previewed nouns). Codability scores were again correlated in Dutch and English ($r=0.49$ for agents in the No Preview condition; $r=0.54$ for events; $r=0.54$ for patients), and did not differ between languages for events and agents ($ps>.3$).

Results

(1) *Structure choice*

Effects of Language, Block, Preview, and Proficiency. Structure choice was not modulated by these variables, either independently or in interactions (all $zs<1.5$; *Figure 1c*).

Effects of Codability. Effects of Codability relative to the experimental manipulations were evaluated in a model testing all two-way interactions between Language, Preview, Event Codability, and Agent Codability. Block was included as an additive factor, and the model included random slopes for all factors except Block and Event Codability. None of the three-way interactions reached significance.

The model showed an interaction between Preview and Agent Codability ($\beta = -.51$, $z = -1.96$, $p = .05$). Easy-to-name agents were more likely to become sentence subjects than harder-to-name agents in the No Preview condition in both L1 and L2. Importantly, there were more active sentences produced in the Preview condition, and the difference between easy-to-name and harder-to-name agents disappeared in this condition. This shows that the Preview manipulation was successful in facilitating agent naming. There was also a weak interaction of Preview with Event Codability ($\beta = -.45$, $z = -1.80$, $p = .07$): high-codability events were described with active syntax in both Preview conditions at very high rates (over 90% actives), and low-codability events were described with active syntax more often in the Preview condition. Together, these effects confirm that agents are more likely to be assigned the subject role when they were easy to encode in L1 as well as L2, both due to their inherent ease of naming and due to the experimental manipulation of naming ease.

(2) *Speech onsets*

Speech onsets were shorter in active than passive sentences in both Dutch (1836 vs. 2152 ms) and English (2005 vs. 2380 ms). Further analyses were run on active sentences.

Effects of Language, Block, Preview, and Proficiency. Speech onsets were shorter in Dutch than in English ($\beta = 173$, $t = 3.07$), and there were no reliable effects of Block, Preview or LexTale scores (*Figure 2c*).

Effects of Codability. Effects of codability were examined in a model including all three-way interactions between Language, Block, Preview, Event Codability, and Agent Codability (with all random slopes). This model showed again that onsets were shorter in Dutch than English. Importantly, onsets were also shorter in items with easy-to-name agents ($\beta = 124$, $t = 4.43$, for the main effect of Agent Codability), and this effect was further modulated by Preview ($\beta = -.63$, $t = -1.93$, $p = .06$): sentences with harder-to-name agents showed greater facilitation in the Preview condition. Finally, there was an interaction between Language, Preview, and Event Codability ($\beta = -142$, $t = -2.32$): speech onsets were longer in sentences describing lower-codability than higher-codability events, but Preview reduced onsets in lower-codability events in English.

(3) *Time-course of formulation*

First Block

Before 400 ms, there were fewer agent-directed fixations overall in the Preview than the No Preview condition in both Dutch and English (main effect of Preview, reliable by participants; Table 5a; *Figure 5a, 5c*). Specifically, there were fewer agent fixations at the beginning of this time window in the Preview than the No Preview condition, but these fixations rose more quickly (by-participant interaction of Preview with the Linear term), and had a smaller and delayed peak (by-participant interaction of Preview with the Quadratic term, by-participant and by-item interactions with the Cubic term; Table 5a).

Importantly, the effect of Language in this time window was similar to that observed in Experiment 1: there were fewer agent fixations in English than in Dutch, and agent fixations dropped in English towards the end of this time window (interaction of Language with the Cubic term).

Table 5.

GCA results for active sentences in Experiment 3. All models include by-participant and by-item random slopes for Language and Preview.

Effect	By-participants			By-items		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
<i>First block</i>						
<i>a) 0-400 ms</i>						
Intercept	-.85	.06	-13.89*	-.64	.07	-8.56*
Linear	5.62	.24	23.81*	3.87	.25	15.22*
Quadratic	-2.00	.16	-12.82*	-1.04	.19	-5.60*
Cubic	-.17	.16	-1.02	-.34	.11	-3.08*
Language	-.14	.12	-1.17	-.12	.06	-1.94*
Preview	-.17	.09	-2.00*	.07	.07	-.97
Language * Preview	-.11	.17	-.66	.13	.03	-4.58*
Linear * Language	.47	.47	1.00	.17	.08	2.05*
Quadratic * Language	.25	.31	.81	.31	.08	3.86*
Cubic * Language	-.65	.33	-1.97*	-.38	.08	-4.89*
Linear * Preview	.37	.11	3.40*	.02	.08	.23
Quadratic * Preview	.30	.10	2.86*	.05	.08	.56
Cubic * Preview	.35	.10	3.69*	.30	.08	3.70*
Linear * Lang * Preview	-.04	.22	-.20	-.22	-.16	-1.34
Quadratic * Lang * Preview	-.55	.21	-2.64*	-.25	-.16	-1.55
Cubic * Lang * Preview	.05	.19	.27	.09	.16	.55
<i>b) 400-2000 ms</i>						
Intercept	.14	.04	3.29*	.15	.05	2.79*
Linear	-2.27	.40	-5.63*	-1.97	.36	-5.47*
Quadratic	-2.49	.31	-7.91*	-2.21	.33	-6.71*

Cubic	1.32	.32	4.11*	1.09	.28	3.92*
Language	.04	.09	.50	1.06	.05	1.27
Preview	.17	.05	-3.12*	-.09	.05	-1.74
Language * Preview	.16	.11	-1.46	-.03	.02	-1.55
Linear * Language	2.13	.80	2.64*	1.45	.09	15.53*
Quadratic * Language	-.21	.63	-.33	-.32	.09	-3.50*
Cubic * Language	-1.29	.64	-2.00*	-.90	.09	-9.72*
Linear * Preview	-1.57	.09	-16.64*	-.65	.10	-6.60*
Quadratic * Preview	.04	.09	.37	-.43	.10	-4.41*
Cubic * Preview	.68	.09	7.23*	.57	.10	5.78*
Linear * Lang * Preview	-1.12	.19	-5.94*	-.67	.19	-3.61*
Quadratic * Lang * Prev.	-.97	.19	-5.15*	-.86	.18	-4.65*
Cubic * Lang * Preview	.08	.19	.44	.21	.19	1.14

*Second block**c) 0-400 ms*

Intercept	-.66	.05	-14.03*	-.54	.07	-7.28*
Linear	4.77	.22	21.33*	3.44	.27	12.82*
Quadratic	-2.06	.16	-13.06*	-1.28	.24	-5.34*
Cubic	.25	.16	1.54	-.12	.12	-1.00
Language	.23	.09	2.50*	.08	.07	1.10
Preview	-.21	.10	-2.24*	-.02	.06	-.40
Language * Preview	-.44	.19	-2.29*	-.32	.03	-10.02*
Linear * Language	-.52	.45	-1.15	-.45	.09	-5.26*
Quadratic * Language	-.06	.32	-.19	-.05	.09	-.63
Cubic * Language	-.14	.32	-.43	-.10	.08	-1.18
Linear * Preview	.14	.11	1.25	.36	.09	4.03*
Quadratic * Preview	.37	.11	3.42*	.06	.09	.71
Cubic * Preview	-.01	.10	-.01	-.15	.09	-1.74
Linear * Lang * Preview	.18	.22	.81	.47	.17	2.76*
Quadratic * Lang * Prev.	.99	.22	4.59*	.69	.17	4.06*
Cubic * Lang * Preview	-.54	.20	-2.71*	-.40	.17	-2.41*

d) 400-2000 ms

Intercept	.25	.04	6.14*	.19	.05	3.56*
Linear	-1.86	.42	-4.41*	-1.43	.47	-3.05*
Quadratic	-2.59	.29	-8.91*	-2.39	.34	-6.96*
Cubic	1.09	.29	3.76*	1.15	.34	3.33*
Language	.23	.08	2.85*	.21	.05	4.58*
Preview	-.16	.06	-2.54*	-.14	.06	-2.46*
Language * Preview	-.03	.13	-.21	-.04	.02	-1.81†
Linear * Language	2.24	.84	2.65*	2.19	.10	22.58*
Quadratic * Language	-.43	.58	-.73	.07	.10	.69
Cubic * Language	-1.00	.58	-1.72†	-.43	.10	-4.41*
Linear * Preview	-.93	.10	-9.02*	-.85	.11	-8.03*
Quadratic * Preview	.08	.10	.74	.40	.11	3.85*
Cubic * Preview	1.20	.10	11.69*	1.00	.11	9.52*
Linear * Lang * Preview	.05	.21	.25	.57	.19	2.96*
Quadratic * Lang * Prev.	-1.47	.21	-7.18*	-.88	.19	-4.52*
Cubic * Lang * Preview	-.36	.21	-1.74†	-.72	.19	-3.73*

* $p < .05$ † $p < .10$

After 400 ms, there was also a strong effect of Preview and Language (Table 5b; *Figure 5a, 5c*). Fixations to previewed agents had a steeper downward slope (interaction between Preview and the Linear term) as well as a faster rise and fall than fixations to non-previewed agents (interaction between Preview and the Cubic term). This suggests that linguistic encoding of the agent was initiated earlier and completed earlier when this character name was familiar.

The effect of Language was again similar to that in Experiment 1. There was a delayed rise in agent fixations in English compared to Dutch, suggesting that speakers began encoding this character linguistically later in L2 than in L1. Towards the end of this time window, speakers also began shifting their attention away from the agent earlier in Dutch than in English, which shows again that linguistic encoding of this character took less time in L1 than L2. In other words, agent fixations had a more shallow downward slope in English (interaction between Language and the Linear term) and were more sustained in English than in Dutch (interaction between Language and the Cubic term).

The effect of Preview was stronger in L2 than in L1. Towards the end of the 400-2000 ms time window, agent fixations dropped more quickly and had a sharper peak in the Preview than No Preview condition, and the magnitude of this effect was larger in English than in Dutch (interaction between Language, Preview, and the Linear and Quadratic terms). Thus by 2000 ms, the distribution of agent fixations in the Preview condition was similar in Dutch and English.

Second Block

As in previous experiments, there was a shift towards linearly incremental planning in Block 2. In the 0-400 ms time window, speakers directed their attention to agents more quickly after picture onset than they did in Block 1. Once again, this shift was stronger in English (i.e., after participants had completed Block 1 in Dutch): fixations to the agent rose and fell more quickly in the No Preview condition in English than in all remaining conditions (interactions of Language, Preview, and the Quadratic and Cubic terms; Table 5c; *Figure 5b, 5d*).

After 400 ms, agent fixations had a steeper downward slope in the Preview than No Preview condition (interaction of Preview with the Linear term; Table 5d; *Figure 5b, 5d*), and rose and fell more quickly in the Preview than No Preview condition (interaction of Preview with the Cubic term). Agent fixations also had a steeper downward slope in Dutch (interaction of Language and the Linear term) and a faster rise and fall in Dutch than in English (interaction of Language and the Cubic term). As in Block 1, these effects show that linguistic encoding of the agent was completed more quickly when speakers used a previewed name and when they prepared to produce L1 sentences.

There were also joint effects of Preview and Language: the distribution of agent fixations had a sharper peak in Dutch previewed sentences (interaction of Language, Preview and the Quadratic term), and the sharpest rise and fall in Dutch previewed sentences compared to all other conditions (interaction of Language, Preview, and the Cubic terms).

Joint analysis of the First and Second Block

The joint analysis of the first and second block confirmed the presence of a shift towards Linear Incrementality across blocks (see Supplementary Material E for the output).

Between 0 and 400 ms, there were more agent-directed fixations in Dutch in Block 1 and in English in Block 2 (interaction of Language and Block). Fixations to the agent rose more quickly in Block 2 (interaction of Block and the Linear term) and were more sustained in Block 2 than Block 1 (interaction of Block and the Cubic term). Interactions of Block with Preview and Language were observed primarily in the by-item analysis. The effect consistent across by-participant and by-item analyses was the interaction of Block, Preview and the Cubic term: agent-directed fixations were more sustained in the No Preview condition of Block 2.

In the 400-2000 ms time window, agent fixations dropped more slowly in Block 2 than Block 1 (interaction of Block with the Linear term). Thus, speakers shifted their attention away from the agent earlier in Block 1, suggesting that they began adding the patient character to the developing sentence earlier in Block 1 than in Block 2. The effect of Preview changed from Block 1 to Block 2: agent fixations dropped more quickly in the Preview condition in Block 1 than Block 2 (interaction between Block, Preview and the Linear term) and had a less sustained peak in Block 1 than Block 2 (interaction between Block, Preview and the Cubic term).

Discussion

Experiment 3 showed the availability of agent names (both the inherent ease of naming and the manipulated ease of naming these characters) influenced structure choice and speech onsets in L1 and L2, as expected, but time-course analyses showed a striking absence of analogous effects in early formulation. Replicating previous research, speakers were more likely to begin sentences with easier-to-name agents than harder-to-name agents, and with previewed than non-previewed agents in L1 and L2 (*Results section 1*). Sentences with easy-to-name agents and with previewed agents also had shorter onsets, demonstrating that the availability of agent names speeded up lexical encoding in both L1 and L2 (*Results section 2*).

More importantly, time-course analyses in Block 1 showed that previewing agent names had little impact on the L1-L2 difference in early eye movements (0-400 ms and the beginning of the 400-2000 ms time window; *Results section 3*). Interestingly, speakers showed a *smaller* preference for the agent in the 0-400 ms time window in the Preview than the No Preview

condition: they chose to encode information about both characters early in the formulation process instead of prioritizing encoding of the subject character. This suggests that sentence planning was not sensitive to the ease of agent naming in the earliest stages. Facilitating agent naming produced a strong effect only during the time window associated with linguistic encoding (400-2000 ms), where agents with previewed names were fixated for less time than agents with non-previewed names.

The effect of Language on formulation was consistent with Experiment 1 and with the No Preview condition of Experiment 2. Speakers were more likely to distribute their attention between the agent and the patient before 400 ms in English than in Dutch, and they also shifted their attention to the agent more slowly in English than in Dutch after 400 ms. This delay in agent-directed fixations in L2 was observed both in the Preview and the No Preview conditions. Since the agent was easy to name and was the first content word of active sentences, speakers could have directed their attention to this character immediately after 400 ms as in Dutch; the fact that they did not do so suggests that the delay in the onset of linguistic encoding cannot be accounted for by difficulties in naming the subject character.

Joint effects of Language and Preview were only observed in the second time window of Block 1 (400-2000 ms): speakers needed less time to encode previewed agents linguistically, and this facilitatory effect of Preview was stronger in English. In other words, increasing familiarity with agent names facilitated processing to a greater extent in the less practiced language (L2) than the native language (L1), but this familiarity benefit was limited to *linguistic encoding*.

In sum, familiarity with agents reduced the time needed to encode their names linguistically but it did not influence the early coordination of message-level and sentence-level processes in L1 and L2: speakers still preferred to delay linguistic encoding in L2 in Block 1. This is the opposite pattern to what one might expect under Linear Incrementality, where properties of the sentence-initial content word would largely determine the distribution of early eye movements. A shift towards linearly incremental planning was observed again in Block 2, replicating previous experiments.

Experiment 4

Experiment 3 showed that early L2 planning was not sensitive to the ease of encoding the sentence-initial content word (i.e., the agent name), suggesting that speakers were less likely to use a linearly incremental planning strategy in L2 than in L1. An alternative explanation for this result is that the delay in the onset of L2 linguistic encoding is influenced by the ease of encoding information that is mentioned *after* the subject character, i.e., the sentence verb. Early

sensitivity to properties of the verb implies sensitivity to the ease of encoding *relational* information, which is more consistent with Hierarchical Incrementality.

To test the extent to which speakers are sensitive to the ease of encoding the sentence verb in L1 and L2, Experiment 4 used a preview manipulation with action verbs. The experiment was again divided into two blocks, and speakers received a preview of the most frequent verbs used for describing half of the target pictures before each block, intermixed with filler verbs.

The Preview manipulation was not expected to have a strong influence on structure choice, although familiarity with individual verbs may allow participants to produce lexically and structurally more complex descriptions. The main hypothesis concerned, again, the time-course of formulation. If speakers delay linguistic encoding of the agent in L2 (Experiments 1-3) because they prefer to encode a larger message first – i.e., a message framework that includes information about the relationship between event characters – then familiarity with suitable verbs should facilitate early planning of relational information (also see [Konopka & Kuchinsky, 2015](#)). This should reduce the L1-L2 difference in early eye movements observed in Experiments 1-3. Dutch and English provide a strong test of this hypothesis because verbs are produced sentence-medially: early effects of verb availability demonstrate extensive planning of the relational content of an event. However, if the reason for the delay in linguistic encoding in L2 lies elsewhere, the Preview manipulation should not change the time-course of formulation and the results should replicate Experiment 3.

Method

Participants

Participants were 62 native Dutch speakers (46 female), with a mean age of 21.4 years ($SD=2.18$, range=18-27; background data were missing from 5 participants), from the same participant pool as in previous experiments. Data from 3 additional participants were excluded due to technical problems.

Table 1 lists descriptive statistics for this sample. LexTale scores were again higher for Dutch than English (Table 2; $t[59]=11.93$; LexTale data were missing from two participants), and were highly correlated. Participants' self-reported English speaking-ability scores were also strongly correlated with their English LexTale scores (Table 2).

Materials, design, and procedure. The materials were the same as in previous experiments. One block was completed in L1 and one in L2. Half of the items from each block were selected to be previewed, and the selected verbs were the most frequent verbs used to describe the event action in previous experiments. The two halves of the experimental items had similar Event and Agent Codability ratings. Counterbalancing the identity of the previewed

verbs, the left-right orientation of the target pictures, block language and block order within participants and within items resulted in 16 experimental lists.

The Preview before each block consisted of verbs suitable for describing 20 of the target events along with 22 randomly chosen verbs. Presentation time were again quite fast (500 ms per word), and participants were asked to read the words to themselves silently. As in Experiment 3, they were told that familiarity with these words would facilitate performance in the main task.

Sentence scoring, codability scoring, and analyses. Applying the same scoring criteria as in previous experiments left 2426 sentences for analysis (86% actives, 12% passives, 2% truncated passives). Scoring and analyses were carried out as in Experiment 3 (*Results sections 1, 2, 3*).

Speakers sometimes used non-previewed verbs on previewed trials (e.g., the non-previewed verb “to bite” to describe the picture of a bee stinging a man) and also used previewed verbs on trials designed to have no preview. Thus, analyses were carried out after recoding all responses into sentences that did and did not use the previewed verbs, irrespective of condition, as in Experiment 3. There were 940 sentences with previewed verbs and 1446 sentences with non-previewed verbs. Again, speakers were as likely to use the previewed agent names in Dutch (39% of 1208 sentences) as they were in English (40% of 1178 sentences).

Codability scores were highly correlated in Dutch and English ($r=0.75$ for agents; $r=0.51$ for events in the No Preview condition; $r=0.65$ for patients). Event and Agent Codability ratings did not differ across languages; Patient Codability ratings were higher in Dutch than in English.

Results

(1) Structure choice

Effects of Language, Block, Preview, and Proficiency. A model including the three-way interaction between Block, Language, Preview and LexTale scores (as well as random slopes for all factors except Preview) showed an interaction between Language, Block and Lextale ($\beta=-1.29$, $z=-2.16$; *Figure 1d*). Speakers produced more active sentences in Block 1, particularly in English, than in Block 2, where structure choices in Dutch and English were comparable. In English, high production of actives in Block 1 was due to lower-proficiency speakers choosing active syntax more often.

Effects of Codability. Effects of codability were evaluated in a model testing all three-way interactions between Language, Block, Preview, Event Codability, and Agent Codability (with random slopes for Language, Block and Preview). The model showed a strong interaction between Preview, Block, and Event Codability ($\beta=-1.52$, $z=-3.78$). Preview modulated structure selection primarily in low-codability events. When using non-previewed verbs, there was little

difference in structure choice for high-codability and low-codability events; when using previewed verbs, speakers produced more active sentences to describe low-codability than high-codability events. These effects were again observed in Block 1 but not Block 2. There were no effects of Agent Codability.

(2) *Speech onsets*

Speech onsets were shorter in active than passive sentences in both Dutch (1746 vs. 1846 ms) and English (2031 vs. 2119 ms). Further analyses were run on active sentences.

Effects of Language, Block, Preview, and Proficiency. Speech onsets were shorter in Dutch than in English ($\beta=290$, $t=6.27$), particularly in speakers with higher LexTale scores ($\beta=-146$, $t=-3.06$, for the interaction of Language with LexTale). LexTale scores were better predictors in Block 1 than Block 2 ($\beta=-99$, $t=-2.14$, for the interaction of Block with LexTale), particularly when speakers completed Block 2 in English (*Figure 2d*).

Onsets were also shorter in sentences with previewed verbs ($\beta=114$, $t=4.18$) in both L1 and L2. This finding suggests that speakers generally preferred to encode information they would mention after the sentence subject (i.e., the verb) in the No Preview condition and that familiarity with a verb in the Preview condition can reduce the preference for early extensive planning.

Effects of Codability. Effects of Codability were examined in a model including all three-way interactions between Language, Block, Preview, Event Codability, and Agent Codability (with all random slopes). This model showed again that onsets were shorter in Dutch than English ($\beta=294$, $t=5.93$), and shorter in items with easy-to-name agents ($\beta=133$, $t=5.31$). Onsets were also shorter in the Preview than the No Preview condition ($\beta=98$, $t=3.73$), and this effect was further modulated by Block and Event Codability ($\beta=-152$, $t=-3.07$, for the three-way interaction). In Block 1, speech onsets were faster for high-codability than low-codability events in the No Preview condition, and this effect was much larger in the Preview condition. Thus, speakers consistently attempted to encode information beyond the sentence subject before speech onset, and the Preview manipulation either facilitated more extensive early planning or it primed speakers to engage only in minimal pre-planning of the verb on the assumption that this word would be easy to produce after speech onset (see time-course results below). Importantly, there were no interactions with Agent Codability, suggesting that the extent to which speakers attended to the verb did not depend on properties of the first content word of the sentence.

(3) *Time-course of formulation*

First Block

Before 400 ms, fixations to agents rose more quickly in the Preview than the No Preview condition (interaction of Preview with the Linear term; Table 6a; *Figure 6a, 6c*). The effect of

Preview depended on Language (interaction of Language, Preview and the Linear term): unlike previous experiments, agent-directed fixations rose more steeply in the Preview condition in English than in Dutch. Thus, by the end of this time window (400 ms), the distribution of agent fixations was similar in the Preview and the No Preview condition in Dutch, but there were more agent fixations in English in the Preview than the No Preview condition. The English distribution was also more peaked (interaction of Language with the Quadratic term).

Table 6.

GCA results for active sentences in Experiment 4. All models included by-participant or by-item slopes for Language and Preview, except models (b) and (d) which only included by-participant slopes for Preview.

Effect	By-participants			By-items		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
<i>First block</i>						
<i>a) 0-400 ms</i>						
Intercept	-.77	.05	-15.22*	-.62	.09	-7.30*
Linear	5.01	.21	23.42*	3.84	.27	14.14*
Quadratic	-2.04	.14	-15.01*	-1.37	.22	-6.16*
Cubic	.15	.15	1.02	-.14	.12	-1.10
Language	.01	.10	.10	.03	.06	.54
Preview	.16	.09	1.75	.12	.06	2.01*
Language * Preview	-.03	.18	-.18	-.04	.02	-1.79
Linear * Language	-.06	.43	-0.15	-.27	.08	-3.49*
Quadratic * Language	-.49	.27	-1.81†	-.49	.08	-6.41*
Cubic * Language	-.04	.30	.12	.02	.07	.32
Linear * Preview	-.44	.10	-4.60*	-.66	.08	-8.51*
Quadratic * Preview	.28	.09	3.04*	.07	.08	.90
Cubic * Preview	.03	.09	.37	.11	.07	1.47
Linear * Lang * Preview	.54	.19	2.85*	.58	.15	3.78*
Quadratic * Lang * Preview	-.10	.18	-.53	-.14	.15	-.94
Cubic * Lang * Preview	-.11	.17	-.63	-.05	.15	-.35
<i>b) 400-2000 ms</i>						
Intercept	.22	.04	6.46*	.19	.05	4.12*
Linear	-2.29	.46	-5.01*	-1.64	.41	-4.02*
Quadratic	-2.83	.31	-9.02*	-2.31	.38	-6.16*
Cubic	1.20	.27	4.41*	.86	.29	2.95*
Language	.16	.07	2.31*	.06	.05	1.22
Preview	.01	.05	.15	-.05	.05	-1.11
Language * Preview	.16	.10	1.59	-.15	.09	-9.90*
Linear * Language	2.21	.91	2.42*	1.52	.09	17.64*
Quadratic * Language	.26	.63	.41	.23	.09	2.70*
Cubic * Language	-.54	.54	-.99	-.29	.09	-3.39*

Linear * Preview	-.31	.09	-3.54*	-.30	.09	-3.46*
Quadratic * Preview	.57	.09	6.43*	.33	.09	3.83*
Cubic * Preview	.11	.09	1.20	.39	.09	4.50*
Linear * Lang * Preview	-.63	.17	-3.63*	-.60	.17	-3.46*
Quadratic * Lang * Prev.	2.04	.18	11.56*	1.87	.17	10.77*
Cubic * Lang * Preview	-.70	.18	-3.98*	-.18	.17	-1.02

*Second block**c) 0-400 ms*

Intercept	-.66	.06	-11.36*	-.56	.09	-6.52*
Linear	4.27	.21	20.59*	3.25	.28	11.71*
Quadratic	-1.67	.18	-9.35*	-1.31	.24	-5.46*
Cubic	.61	.12	5.15*	.45	.13	3.37*
Language	.05	.12	.41	-.08	.05	-1.52
Preview	.10	.08	1.26	.07	.07	.96
Language * Preview	.04	.16	.23	.03	.03	1.23
Linear * Language	.45	.41	1.09	.40	.08	5.18*
Quadratic * Language	.06	.36	.16	.14	.08	1.86*
Cubic * Language	.02	.24	.11	-.03	.08	-.44
Linear * Preview	-.25	.09	-2.69*	-.13	.08	-1.62
Quadratic * Preview	.27	.09	2.97*	.34	.08	4.32*
Cubic * Preview	-.13	.08	-1.54	-.16	.08	-2.05*
Linear * Lang * Preview	.42	.18	2.28*	.42	.16	2.69*
Quadratic * Lang * Prev.	1.00	.18	5.49*	.98	.16	6.29*
Cubic * Lang * Preview	-.13	.17	-.78	-.11	.16	-.71

d) 400-2000 ms

Intercept	.16	.04	4.47*	.14	.05	2.82*
Linear	-2.38	.38	-6.18*	-1.83	.45	-4.09*
Quadratic	-1.91	.31	-6.23*	-1.69	.34	-4.91*
Cubic	1.03	.25	4.13*	1.03	.31	3.29*
Language	.15	.07	2.08*	.10	.04	2.48*
Preview	-.05	.05	-.96	-.03	.03	-.97
Language * Preview	.01	.11	.07	.03	.02	2.12*
Linear * Language	.37	.77	.49	.70	.09	7.75*
Quadratic * Language	1.13	.61	1.85†	1.16	.09	12.91*
Cubic * Language	-.67	.50	-1.35	-.51	.09	-5.74*
Linear * Preview	.09	.09	1.02	-.12	.09	-1.28
Quadratic * Preview	.29	.09	3.14*	.41	.09	4.29*
Cubic * Preview	-.49	.09	-5.26*	-.48	.09	-5.19*
Linear * Lang * Preview	-1.52	.19	-8.22*	-2.14	.18	-11.75*
Quadratic * Lang * Preview	-.02	.19	-.10	-.18	.18	-1.02
Cubic * Lang * Preview	-.41	.19	-2.23*	.31	.18	1.72

* $p < .05$ † $p < .10$

Between 400 and 2000 ms, eye movements were sensitive to both Preview and Language (Table 6b; Figure 6a, 6c). As in previous experiments, speakers fixated the agent preferentially after 400 ms and started shifting their attention away from the agent in the second half of this window. Compared to Experiments 1-3, however, the shift of gaze to the agent after 400 ms was delayed, suggesting that speakers generally initiated linguistic encoding later in Experiment 4.

This is likely a result of the Preview manipulation, which drew attention to encoding of the verb rather than encoding of the first content word (the agent).

Importantly, the effect of Preview was different in L1 and L2. By 2000 ms, there was a steeper drop in agent-directed fixations in Dutch sentences than in English sentences (interaction of Language and the Linear term), because speakers completed linguistic encoding of the agent more quickly in L1 than L2. In Dutch, speakers also showed a stronger preference to fixate the agent in the Preview than the No Preview condition. In English, however, there were more agent fixations in the No Preview condition, while the distribution of agent fixations in the Preview condition was more flat (interactions of Preview, Language and the Linear and Quadratic terms): speakers showed a smaller preference for the agent but maintained their attention on this character for a longer time in English than in Dutch. Thus, in the Preview condition in English, speakers devoted more attention to the agent before 400 ms and less attention after 400 ms.

Second Block

Before 400 ms, fixations to the agent in the Preview condition rose more quickly in English sentences than Dutch sentences (interaction of Language, Preview, and the Linear term; Table 6c, *Figure 6b, 6d*). The distribution of agent fixations in the Preview condition was also more flat in English than Dutch sentences (interaction of Language, Preview, and the Quadratic term). Thus, at 400 ms, speakers were more likely to fixate agents in English (i.e., after completing Block 1 in Dutch) than in Dutch (i.e., after completing Block 2 in English). This rapid increase in agent-directed fixations in English is consistent with previous experiments, and shows more linearly incremental planning in Block 2 in English.

Importantly, after 400 ms, speakers spent less time fixating agents in the Preview condition than the No Preview condition in both L1 and L2: agent-directed fixations had a sharper peak in the Preview condition (interaction of Preview with the Quadratic term) and a sharper drop towards the end of this time window in the Preview than No Preview condition (interaction of Preview with the Cubic term; Table 6d, *Figure 6b, 6d*). Thus, speakers began directing their attention away from the agent earlier when they used a previewed verb. Speakers also spent more time fixating agents in English than in Dutch, and the English distribution was more flat than the Dutch distribution (interaction of Language with the Quadratic term). Finally, speakers showed the fastest decline in agent-directed fixations in Dutch sentences and the slowest decline in English sentences with non-previewed verbs (interaction of Language, Preview, and the Linear term), showing late addition of the patient to the sentence when the verb was not easy to encode in English.

Joint analysis of the First and Second Block

The joint analysis of the first and second block showed a different pattern than in earlier experiments (see Supplementary Material F). In the 0-400 ms window, fixations to the agent rose more slowly in Block 2 (interaction of Block with the Linear term) and were less variable in Block 2 than Block 1 (interaction of Block with the Cubic term), irrespective of Language. In other words, speakers were more likely to quickly encode the agent in Block 1 than Block 2.

Agent-directed fixations in the 400-2000 ms window had a more flat peak in Block 2 than Block 1 (interaction of Block with the Quadratic term), as speakers fixated the agent longer in Block 2. The model also showed that agent-directed fixations dropped most quickly in the Preview condition in Block 1 (interactions of Block, Preview, and the Linear term). Agent fixations showed the most fluctuations in the Preview condition in Block 1 and the least fluctuations in Block 2 (interactions of Block, Preview, and the Cubic term). In other words, speakers shifted their attention towards the agent and, at the end of this time window, towards the patient earlier in Block 1 than Block 2.

Discussion

Experiment 4 showed that manipulating verb familiarity had a small influence on structure choice, and that the effects of Preview depended on properties of the events (Event Codability) but not properties of the event agents (Agent Codability; *Results section 1*). Speakers with lower LexTale scores were again more likely to use active syntax in English, particularly when describing lower-codability events. The Preview manipulation had a stronger effect on speech onsets (*Results section 2*): speakers began sentences with previewed verbs more quickly than sentences with non-previewed verbs, and these effects were also modulated by Event Codability but not Agent Codability. This Preview effect suggests that early sentence planning had a fairly extensive scope as it included the sentence verb. It is of course possible that this was due to the Preview manipulation drawing attention to the verb. Nevertheless, the robustness of this effect in both Dutch and English, as well as the absence of any effects of Agent Codability, suggests greater sensitivity to relational than non-relational information during planning.

Importantly, the verb Preview manipulation had a different effect on sentence planning than the noun Preview manipulation in Experiment 3 (*Results section 3*). Whereas speakers' familiarity with agent names did not influence early planning and thus did not eliminate the early L1-L2 difference in eye movements in Experiment 3, familiarity with verbs in the current experiment increased the likelihood of speakers fixating *agents* early in the planning process in L2. Thus, compared against Experiment 3, the peak of the English distribution in the second time window (400-2000 ms) was not delayed relative to the Dutch distribution: facilitating encoding of the verb reduced the need for extensive early planning of the message and, by extension,

reduced the delay in the onset of linguistic encoding in L2. These results suggest that early sentence planning in L2 is particularly sensitive to information that is expressed linguistically with the sentence verb. In the second time window, there were also fewer agent-directed fixations in the Preview condition in L2 than in L1 before speech onset, suggesting that speakers were ready to begin adding the patient to the developing sentence earlier in L2 and thus that the Preview manipulation had a larger benefit in L2 than L1.

An alternative explanation for these results is again that the Preview manipulation explicitly drew attention to verbs, and thus that sentence planning is not normally as sensitive to relational information. At the same time, however, the key observation in this experiment is the presence of an effect of verb preview at all. Compared against Experiment 3, the results show that formulation was more sensitive to the second content word than to the first content word of the sentence. In principle, if speakers prefer to delay linguistic encoding in L2, then they should be more likely to delay planning of the *second* content word than to delay planning of the *first* word, and they should show sensitivity primarily to the properties of the first word.

A related question is whether sensitivity to the sentence verb need imply that the results are consistent with Hierarchical Incrementality. Speakers may, for example, simply show a preference for encoding more than one content word early in the planning process. Testing whether the information value of the second content word is relevant for planning (i.e., whether this word expresses relational or non-relational information) requires a comparison with a language where the main verb is not produced in sentence-medial position (e.g., a language where verbs occur in sentence-initial or sentence-final position; [Hwang & Kaiser, 2014, 2015](#)). Nevertheless, there are reasons to believe that the results do support Hierarchical Incrementality in L2 planning. In particular, Dutch and English are both SVO languages: speakers tend to encode information in an event in the order of mention, and indeed speakers show an increase in agent fixations within 600 ms of picture onset. This leaves relatively little time for properties of the verb to influence early formulation; thus, the fact that sentence planning *is* influenced by verb familiarity at all is strong evidence of early sensitivity to relational information.

Why does familiarity with verbs influence early planning? We propose that this effect occurs because verbs carry information that is compatible with the process that speakers appear to prioritize early during L2 formulation, i.e., *message-level relational* encoding. Verbs express relationships between referents, so familiarity with a verb can decrease the costs of encoding a relational framework for the message and sentence. In other words, it is possible that speakers need less time to encode the preverbal message in the Preview condition because the previewed verbs function as *primes* that increase the efficiency of early conceptual encoding. Importantly,

speakers were more likely to adapt their planning strategies in L2 than in L1 in this experiment, suggesting a broader scope of planning in L2 than in L1.

General Discussion

Summary

Four experiments compared the spontaneous production of simple event descriptions in Dutch (L1) and English (L2), and together, the results showed systematic L1-L2 differences in sentence formulation and accounted for the source of these differences. Participants' production patterns (structure choice and speech onsets) showed departures from L1 production patterns when using their less practiced L2 (English). Then, analyses of the time-course of formulation showed systematic L1-L2 differences that are consistent with the predictions of Linear Incrementality for L1 and Hierarchical Incrementality for L2.

To summarize the results for structure choice (*Results section 1* in all Experiments), the target events elicited a large majority of active descriptions but structure choice was sensitive to a number of speaker-specific and item-specific variables. First, speakers showed more syntactic flexibility in L1: they produced more active sentences in their less practiced L2 and more passive sentences in their highly practiced L1. Second, this L1-L2 difference was mediated by speakers' L1 and L2 vocabulary knowledge: speakers with higher LexTale scores showed more flexibility. Thus, while L1 speakers selected sentence structures on the basis of multiple properties of the target events, L2 speakers were more likely to use active syntax "by default".

To summarize the speech onset results (*Results section 2* in all Experiments), production was faster in L1 than L2 (e.g., [Sadri Mirdamadi & de Jong, 2015](#)). Speech onsets were again sensitive to speakers' vocabulary knowledge and to properties of the events. Since speech onsets are naturally more directly influenced by the ease of retrieving the first content word, this measure did show sensitivity to properties of the agents (Agent Codability) in both L1 and L2 in all but one experiment (Experiment 4).

Our main focus was on L1-L2 differences in the time-course of sentence formulation (*Results section 3* in all Experiments). The global pattern of formulation was similar to previous studies (beginning with [Griffin & Bock, 2000](#)), but speakers consistently employed different planning strategies when producing sentences in a highly practiced (L1) and a less practiced (L2) language. Experiment 1 showed that speakers were more likely to engage in extensive message planning and thus to delay the onset of linguistic encoding in L2 than L1. In other words, planning was more *linearly incremental* in L1 and more *hierarchically incremental* in L2. Experiment 2 verified whether this early L1-L2 difference was indeed driven by speakers' preference to encode a larger message in L2 by comparing planning of descriptions of familiar

and unfamiliar events. The results confirmed that when speakers were familiar with the events and thus needed less time to generate the preverbal message, the L1-L2 difference in early formulation disappeared. Experiments 3 and 4 tested whether speakers choose to delay linguistic encoding in L2 because of difficulty encoding the first content word (the agent name, i.e., non-relational information) or the second content word (the sentence verb, i.e., relational information). Early L2 formulation did not change with increased familiarity of agent names but was sensitive to the manipulation of verb familiarity. This outcome is consistent with L2 planning being more hierarchical than L1 planning.

In sum, the results support a processing model in which both linearly and hierarchically incremental planning are possible (e.g., [Konopka & Meyer, 2014](#).) Earlier studies demonstrated that planning strategies in L1 can be influenced on a moment-to-moment basis by the efficiency of individual linguistic encoding operations, and the present results show that this conclusion extends to L1 and L2 as a whole – i.e., to speakers with higher and lower levels of *overall* experience using a language.

Why is L2 planning more hierarchical?

Both Linear and Hierarchical Incrementality have advantages and disadvantages for planning, and speakers' engagement in one type of planning over the other can reflect an implicit cost-benefit assessment carried out at the moment of speaking. In the case of L2 planning, the benefits of early preparation may outweigh the costs: speakers may prefer extensive, early planning in L2 because it allows them to prepare a larger message and to begin linguistic encoding only once they know what they want to say. This reduces processing load after 400 ms, as it leaves primarily linguistic encoding to be completed.

Previous research comparing offline production of L1 and L2 sentences has also shown that speakers benefit from receiving some preparation time when producing L2 utterances ([Crookes, 1989](#); [Ellis & Yuan, 2004](#); [Foster & Skehan, 1996](#)). L2 speakers and writers are more likely to produce longer, structurally complex, fluent and error-free utterances, and to use more discourse markers when given preparation time (counted in *minutes*) prior to beginning overt production. Here we used a paradigm that provides a closer analogue to production in every-day situations (and thus we measure planning on a *millisecond* scale), but similar processes appear to be at work when generating sentences spontaneously and semi-spontaneously across studies. More specifically, we show that speakers use the first few hundred milliseconds after trial onset as conceptual “preparation” time in L2. Of course, generating a message-level framework early on does not (and cannot) guarantee that speakers will not run into difficulties when they begin linguistic encoding: speakers may still experience production delays if they have trouble

retrieving individual words after 400 ms. Nevertheless, the results show that speakers choose to distribute their resources by temporally separating message-level and sentence-level planning and thus engage in more extensive message preparation in L2 than in L1.

Proficiency

In this paper, we describe the difference between L1 and L2 processing as being a question of *experience*. We intentionally use a vague term (“experience”) to refer to differences in speakers’ overall level of practice using a given language: all experiments included only a simple assessment of proficiency, and indeed, part of the recurring challenge in the measurement of L2 proficiency is the identification of the knowledge and skills that allow a speaker to “use” a second language effectively.

Our results showed modulation of structure choice and speech onsets by LexTale scores, demonstrating that L1-L2 differences in these outcome variables can reflect between-speaker differences in explicit linguistic knowledge. Naturally, vocabulary knowledge is a simple, unidimensional index of proficiency; however, increases in linguistic knowledge in L2 learners go hand-in-hand with increased use of the language and thus the development of further language skills. From a functional perspective, a key property of language use in every-day interactions is the ability to express one’s thoughts quickly and fluently (i.e., to be “communicatively efficient”; [de Jong et al., 2012](#)), and our results suggest that mastery of the L2 lexicon can predict L1-like language use. However, further research is necessary to identify the roles that other predictors of L2 production fluency (such as acquisition history) might play in the planning process (see [de Jong et al., 2013](#)).

For exploratory purposes, we also carried out by-participant time-course analyses on English sentences with English LexTale scores as a continuous predictor in each experiment. If proficiency modulates the extent to which L2 speakers show a native-like pattern of sentence formulation, then speakers with higher L2 scores should show similar L1 and L2 formulation patterns. There was support for this hypothesis in all four experiments in the analyses of early eye movements (i.e., in the 0-400 ms time window). In Experiment 1, speakers with higher LexTale scores showed more fixations directed to the agent (consistent with the L1 results) than speakers with lower scores ($t > 2$ for the effect of LexTale scores). In Experiment 2, there were fewer fixations to the agent in the Preview condition and more fixations in the No Preview condition in L1, and again speakers with higher L2 LexTale scores showed similar effects in L2 ($t > 2$ for the interactions of LexTale with Preview and the Linear, Quadratic, and Cubic terms). These between-speaker differences are broadly consistent with the between-speaker differences

observed in Block 1 in Dutch and English: the more proficient L2 speakers showed L1-like patterns of planning.

The results were different in Experiments 3 and 4, where the Preview manipulations consisted of presentation of specific lexical items ($ts > 2$ for the interactions between LexTale scores, Preview and the three time terms in the 0-400 ms time window). In Experiment 3, speakers generally showed little influence of the agent name preview on formulation in L2, suggesting that L2 planning is less driven by the availability of agent names than L1 planning, and L2 speakers with higher LexTale scores in particular were *less likely* to prioritize encoding of the agent. Thus, increasing familiarity of agent names allowed speakers with higher L2 LexTale scores to quickly devote resources to processing of event information produced beyond the first content word of the sentence early on in the planning process. In Experiment 4, the L2 speakers as a group showed a strong influence of verb preview on formulation, suggesting that L2 planning is generally more sensitive to the availability of relational information than L1 planning. Here, speakers with higher L2 LexTale scores in particular were *more likely* to prioritize encoding of the agent. Thus, speakers with higher L2 vocabulary scores showed the least sensitivity to a linguistic manipulation of the ease of encoding non-relational information and the most sensitivity to a linguistic manipulation of the ease of encoding relational information.

Finally, we note that our observations of between-speaker differences are based on performance of students who were already highly proficient in English and had above-average performance on a test of English vocabulary. It is possible that less proficient language learners plan utterances differently, and thus that Linear and Hierarchical Incrementality are not both plausible planning strategies for all learners. For example, beginners may be less concerned with their ability to assemble words into longer utterances and more concerned with their ability to find contextually appropriate words in the first place. Questions about the coordination of message-level and sentence-level encoding processes naturally become an object of study only once speakers have adequate knowledge and experience to use their L2 for communication. Since sentence planning requires rapid coordination of processes that rely on limited resources (Swets & Ferreira, 2002; Wagner et al., 2010), a comprehensive account of L2 planning will likely also require an assessment of working memory and executive control (e.g., see Myachykov et al., 2011).

Task effects

All experiments also provide an example of flexibility in planning within the same task. Changes in planning strategies from Block 1 to Block 2 showed that, whether speakers used their

L1 or L2, sentence formulation became more linear over the course of the experiments: as speakers gained experience with the task, they were more willing to start linguistic encoding of the agent immediately after picture onset in Block 2. This confirms the hypothesis that planning choices are driven by experience – albeit on a much shorter time-scale than language use outside of the lab – and shows that, broadly speaking, the nature of changes in L2 planning strategies are no different from the strategic choices that speakers make when preparing sentences in L1. Thus, our results support the claim that L1-L2 differences are primarily a question of degree in a system that allows a range of planning strategies for both L1 and L2, and that L2 planning can become L1-like.

It is difficult to determine if the shift from more hierarchical planning in Block 1 to more linear planning in Block 2 was stronger in English or in Dutch without additional controls. For example, one hypothesis is that L2 speakers should be more influenced by recent experience than L1 speakers because they have less overall experience with the target language and thus more room for learning. Consequently, each use of a word or structure in L2 should have a stronger influence on the ease of subsequent production than in L1. Nevertheless, an interesting observation is that planning strategies “transferred” from one language to another in the current experiments. When speakers switched from Dutch in Block 1 to English in Block 2, there was an effect of Block 1 planning strategies on Block 2 planning strategies. In other words, speakers became more linearly incremental in L2 when they had previous experience with the task (i.e., Block 1) and not previous experience using their L2 in the same task. This suggests that the coordination of message-level and sentence-level processes is partly modulated by language-specific experience and partly by a non-language-specific tuning of the efficiency of processes at the interface of thinking and speaking.

Flexibility and boundaries in sentence planning

The L1-L2 comparison also raises questions about potential boundaries to the degree of flexibility in sentence planning. Comprehension and production studies have repeatedly shown that speakers and listeners engage in “good enough” processing, and this observation can also be extended to “good enough” planning (Swets, Jacovina, & Gerrig, 2013). Certainly, our results show that speakers can flexibly adapt their planning to the current production context, and indeed any experience-driven account is based on the assumption that encoding strategies can change to support efficient processing. The presence of shifts from one strategy to another suggests that the choice to engage in either linearly and hierarchically incremental planning in L2 is not a result of architectural constraints. Given the range of contexts in which production normally occurs (e.g., settings with different task goals, conversational partners, and communicative

pressures) and the typological diversity of existing languages, flexibility may be among the most useful design features of language.

At the same time, it is likely that there are boundaries to experience-driven flexibility in planning. A central question for planning theories concerns the nature of the linguistic processes that exert the largest influence on planning. One distinction relevant for sentence production is between processes responsible for *relational* encoding (such as syntax) and processes responsible for *non-relational* encoding (such as lexical retrieval). Recent studies suggest that relational processes may receive priority during production (Allum & Wheeldon, 2007, 2009; Konopka, 2012; Konopka & Meyer, 2014; Kuchinsky, 2009). For example, Konopka and Meyer (2014) compared the formulation of descriptions for events that varied in the ease of gist encoding (a relational variable) and events that varied in the ease of character encoding (a non-relational variable). Their results suggested that the ease of gist encoding modulated the extent to which the ease of character encoding also influenced planning, and not the other way around (also see Konopka, 2012). This conclusion was based on a between-item comparison, but it suggests the existence of “soft” boundaries on the range of encoding strategies supported by the language system, and it is consistent with theories giving relational processes (i.e., grammar) a key role in L1 production (Bock, 1987).

If speakers rely on the same architecture to produce sentences in L1 and L2, then group differences can point to theoretically important bottlenecks. The fact that L2 planning was more hierarchically incremental suggests that early generation of a larger message may be the optimal strategy in the weaker language. Further, the comparison between Experiments 3 and 4 (familiarity with nouns vs. verbs) highlights differential sensitivity to relational information in L1 and L2 planning.

Nouns and verbs

Differences in the nature and ease of retrieving nouns and verbs have direct implications for sentence planning. Nouns and verbs differ in their relational properties (verbs express complex relational meanings) and thus in the extent to which they require encoding of other contextually important information. The processing load of noun and verb encoding also differs: verbs are harder to learn than nouns during first language acquisition (Bornstein et al., 2004; Golinkoff & Hirsh-Pasek, 2008; Hirsh-Pasek & Golinkoff, 2010; Mätzig, Druks, Masterson, & Vigliocco, 2009) and they continue to present difficulties in second language acquisition. Thus, arguably, verb familiarity can be expected to have a stronger influence on any language processing task than noun familiarity, and a stronger influence in L2 (the less practiced language) than L1 (the highly practiced language).

A complete investigation of the role of different word classes (nouns vs. verbs) will inevitably include cross-linguistic comparisons of sentence planning. For example, assessing the role of verbs in sentence planning will require comparing planning patterns in languages where verbs occur in sentence-initial, sentence-medial and sentence-final position. As outlined earlier, however, SVO languages like English and Dutch provide a strong test of the hypothesis that encoding of verbs (or conceptual information that is expressed linguistically with a verb) is important during the earliest stages of formulation: verbs occur in sentence-medial position, so speakers do have the option of postponing encoding of the verb until after they have encoded the sentence subject. Speakers are undoubtedly aware (at least occasionally) of the fluency and communicative adequacy of their utterances. Earlier experience using their L1 and L2 may have taught them that encoding relational information is often taxing and, in the absence of strong pressure to begin speaking quickly, should be performed before speech onset in order to avoid disfluencies after speaking has already begun. In this sense, the largest difference between L1 and L2 users, as well as between more proficient and less proficient L2 users, may be defined in terms of speakers' ability to perform relational encoding quickly.

Limitations

SVO sentences are relatively short: they consist of three content words (S, V, O). While studying the coordination of processes needed to encode three concepts and lexical items in an active sentence can be experimentally challenging, these sentences are still much simpler than the range of utterances that speakers produce outside of the lab. For example, the agent and patient names in these studies were simple noun phrases (e.g., *the mailman*) and did not require retrieval of other content words (e.g., *the tall mailman wearing a blue uniform*). Future studies will need to verify the degree of flexibility in planning during production of longer and structurally more complex sentences. If speakers can flexibly allocate resources to different production processes from the outset of formulation, we predict that producing semantically richer descriptions should result in shifts in processing resources that are observably larger in L2 than in L1.

Another place where L1 and L2 formulation may differ is in the timing of speech onsets relative to the timing of encoding the first character. Specifically, speakers may initiate overt production immediately after retrieving the first character name or only after beginning retrieval of the verb. Visual comparison of the plots for L1 and L2 across experiments suggests that speakers may use a different criterion to decide when to initiate speech in L1 and L2 (speech onsets in L1 occur after speakers have already shifted their gaze to the patient). These differences may or may not depend on the planning strategies adopted at picture onset, so this remains to be

established in future research (at present, Linear and Hierarchical Incrementality do not make predictions about the timing of speech onsets).

Finally, an important limitation of any cross-linguistic study are potential differences in L1-L2 lexicons and grammatical systems. Here, we compare two languages that are highly similar in terms of their grammar of transitive sentences, and this allows us to attribute differences in the time-course of L1 and L2 formulation to variables other than language-specific encoding requirements. However, languages vary considerably in their grammatical requirements, so L2 speakers may need to engage in more or less planning of information beyond the first content word of an utterance in different languages. In this regard, an interesting property of Dutch is flexibility in verb placement: verbs are produced in sentence-medial position in SVO sentences, but can also be produced sentence-finally (“De man is het kindje aan het redden”; also see [Konopka & Kuchinsky, 2015](#), for examples of sentences where verbs are produced before character names, e.g., “Vandaag reedt de man het kindje”). The number of sentences with this structure was too small for analysis in the current experiments, so more research is necessary to test for structure-specific planning patterns in Dutch. In particular, the extent to which a speaker’s native language might modulate planning in a second language remains an important question for further study (e.g., see [von Stutterheim et al., 2012, 2013](#)).

Conclusions

The production system provides speakers with a high degree of flexibility in sentence planning, and speakers appear to coordinate message-level and sentence-level processes differently during spontaneous L1 and L2 sentence production. Given the larger processing demands of L2 production, speakers engage in more extensive message-level relational encoding early in the L2 planning process compared to L1 planning. Comparisons between and within speakers suggest that these effects are experience-driven, and thus that the entire continuum of planning strategies may be available to both L1 and proficient L2 speakers.

Acknowledgements

Many thanks to Annelies van Wijngarden, Caitlin Decupyer, and student assistants in the Psychology of Language Department (Esther Kroese, Marloes Gauwamans, Jessica Aguilar Diaz, Ilse Wagemakers) for invaluable help during data collection and processing.

References

- Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 791-810.
- Allum, P. H., & Wheeldon, L. (2009). Scope of lexical access in spoken sentence production: Implications for the conceptual–syntactic interface. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1240-1255.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Languages*, 59, 390-412.
- Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2009). Persistence of emphasis in language production: A cross-linguistic approach. *Cognition*, 112, 300-317.
- Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition*, 127, 287-306.
- Bock, J. K. (1986a). Syntactic persistence in language production. *Cognitive psychology*, 18, 355-387.
- Bock, J. K. (1986b). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 575-586.
- Bock, K. (1987). Exploring levels of processing in sentence production. *Natural Language Generation*, 351-63.
- Bock, K., & Ferreira, V. S. (2014). Syntactically speaking. In *The Oxford handbook of language production*, 21-46. Oxford University Press.
- Bock, J. K., Irwin, D. E., & Davidson, D. J. J. (2004). Putting first things first. In F. Ferreira & M. Henderson (Eds.), *The integration of language, vision and action: Eye movements and the visual world* (pp. 249–278). New York: Psychology Press.
- Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, 75, 1115-1139.
- Bosker, H. R., Pinget, A.-F., Quené, H., Sanders, T. J. M., & De Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30, 159–175.
- Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, 118, 172-189.

- Chabal, S., & Marian, V. (2015). Speakers of different languages process the visual world differently. *Journal of Experimental Psychology: General*, 144, 539-550.
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological review*, 113, 234-272.
- Costa, A. (2005). Lexical Access in Bilingual Production In J. F. Kroll and A. M. B. De Groot, *Handbook of Bilingualism* (pp. 308-325). New York, New York: Oxford University Press, Inc.
- Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11, 367-383.
- Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, 22, 1630-1645.
- De Jong, N. H., Steinel, M., Florijn, A. F., Schoonen, J. J. M., & Hulstijn, J. H. (2008). The effect of task complexity on fluency and functional adequacy of speaking performance. In *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching*. Brussel: Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten.
- De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition*, 34, 5-34.
- De Jong, N. H., Steinel, M. P., Florijn, A., Schoonen, R., & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 34, 893-916.
- Edwards, J. V. (2004). Foundations of Bilingualism In T. K. Bahtia and W. C. Richtie, *The Handbook of Bilingualism* (pp. 7-31). Malden, MA, Blackwell Publishing Ltd.
- Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, 54, 227-275.
- Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26, 59-84.
- Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33, 715-736.
- Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46, 57-84.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second language acquisition*, 18, 299-323.

- Ganushchak, L., Konopka, A. E., & Chen, Y. (2016). Accessibility of referent information influences sentence planning: an eye-tracking study. *Frontiers in Psychology*, 8.
- Glanemann, R., Zwitserlood, P., Bölte, J., & Dobel, C. (2016). Rapid apprehension of the coherence of action scenes. *Psychonomic Bulletin & Review*, 23, 1566-1575.
- Gleitman, L., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, 57, 544–569.
- Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences*, 12, 397-403.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67-81.
- Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, 82, B1–B14.
- Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11, 274-279.
- Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General*, 142, 880.
- Hartsuiker, R. J., & Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychologica*, 128, 479-489.
- Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, 15, 409-414.
- Hirsh-Pasek, K., & Golinkoff, R. M. (Eds.). (2010). *Action meets word: How children learn verbs*. Oxford University Press.
- Hwang, H., & Kaiser, E. (2014). The role of the verb in grammatical function assignment in English and Korean. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 1363-1376.
- Hwang, H., & Kaiser, E. (2015). Accessibility effects on production vary cross-linguistically: Evidence from English and Korean. *Journal of Memory and Language*, 84, 190-204.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformations or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434–446.

- Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, 66, 143–162.
- Konopka, A. E. (2014). *Discourse changes the timecourse of sentence formulation*. Poster presented at the 19th Architectures and Mechanisms for Language Processing Conference [AMLaP], Marseille, France.
- Konopka, A. E., & Brown-Schmidt, S. (2014). Message encoding. *The Oxford handbook of Language Production*, 3-20. Oxford University Press.
- Konopka, A. E., & Kuchinsky, S. E. (2015). How message similarity shapes the time-course of sentence formulation. *Journal of Memory and Language*, 84, 1-23.
- Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, 73, 1–40.
- Kroll, J., & De Groot, A. M. B. (2009). *Handbook of Bilingualism: Psycholinguistic approaches*. New York, New York: Oxford University Press, Inc.
- Kuchinsky, S. E., & Bock, K. (2010). From seeing to saying: Perceiving, planning, producing. *Paper presented at the 23rd CUNY human sentence processing conference, New York*.
- La Heij, W. (2005). *Selection process in monolingual and bilingual lexical access*. In J. F. Kroll and A. M. B. De Groot, *Handbook of Bilingualism* (pp. 308-325). New York, New York: Oxford University Press, Inc.
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44, 325-343.
- Levelt, W. (1989). *Speaking: From intention to articulation*. MA: The MIT Press, Cambridge.
- Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, 114, 29-41.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 940-967.
- Mätzig, S., Druks, J., Masterson, J., & Vigliocco, G. (2009). Noun and verb differences in picture naming: past studies and new evidence. *Cortex*, 45, 738-758.
- Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, 66, B25–B33.
- Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: Chapman & Hall/CRC.

- Myachykov, A., Thompson, D., Scheepers, C., & Garrod, S. (2011). Visual attention and structural choice in sentence production across languages. *Language and Linguistics Compass*, 5, 95-107.
- Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. C. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*, 30, 1187-1208.
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134, 427-459.
- Sadri Mirdamadi, F., & De Jong, N. H. (2015). The effect of syntactic complexity on fluency: Comparing actives and passives in L1 and L2 speech. *Second Language Research*, 31, 105-116.
- Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin Jr, R. D., & Levinson, S. C. (2013, March). Planning units in Tagalog sentence production: Evidence from eye tracking. *26th Annual CUNY Conference on Human Sentence Processing*.
- Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14, 357-385.
- Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56, 153-171.
- Segaert, K., Wheeldon, L., & Hagoort, P. (2016). Unifying structural priming effects on syntactic choices and timing of sentence generation. *Journal of Memory and Language*, 91, 59-80.
- Segalowitz, N. (2010). Cognitive bases of second language fluency. New York: Routledge.
- Smith, M., & Wheeldon, L. (2001). Syntactic priming in spoken sentence production—an online study. *Cognition*, 78, 123-164.
- Von Stutterheim, C., Flecken, M., & Carroll, M. (2013). Introduction: Conceptualizing in a second language. *International Review of Applied Linguistics in Language Teaching*, 51, 77-85.
- Von Stutterheim, C., Andermann, M., Carroll, M., Flecken, M., & Schmiedtová, B. (2012). How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics*, 50, 833-867.

- Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 423–440.
- Zareva, A., Schwanenflugel, P., & Nikolova, Y. (2005). Relationship between lexical competence and language proficiency: Variable sensitivity. *Studies in Second Language Acquisition*, 27, 567-595.

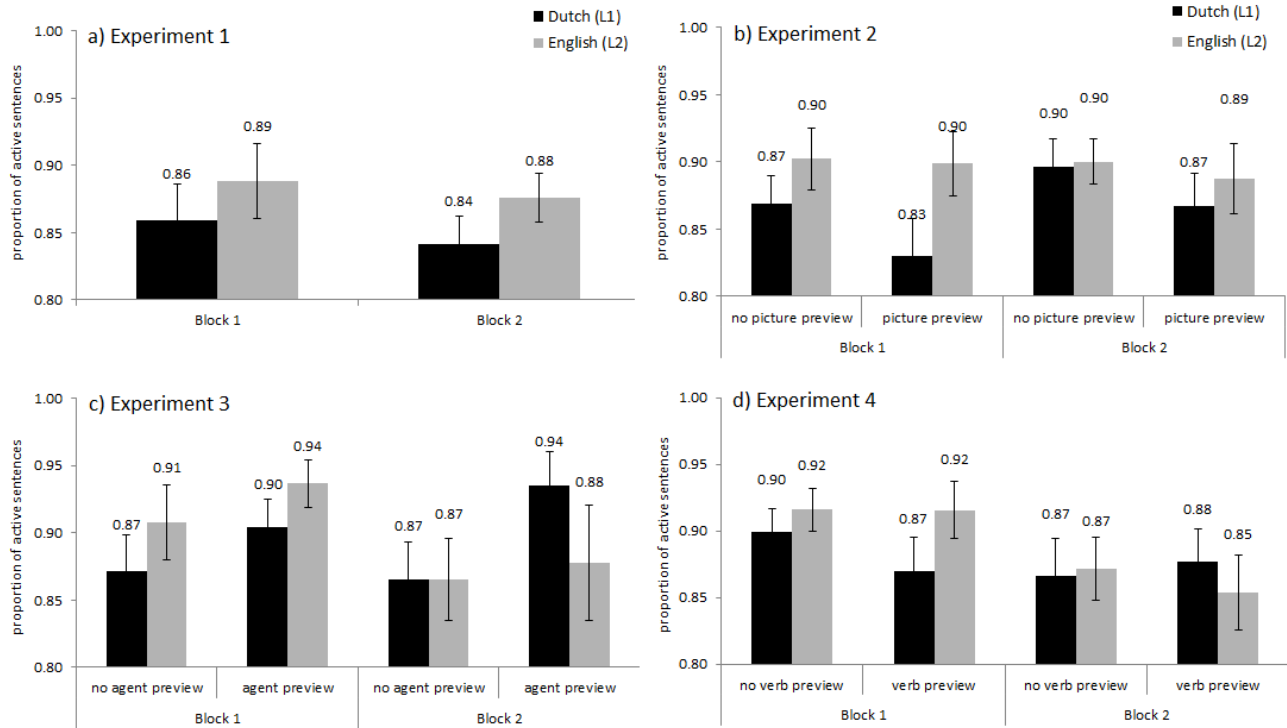


Figure 1. Proportion of active sentences (by-participant means with standard errors) produced as a function of (a) Language and Block in Experiment 1, and (b-d) Language, Block and Preview in Experiments 2-4. Error bars are standard errors.

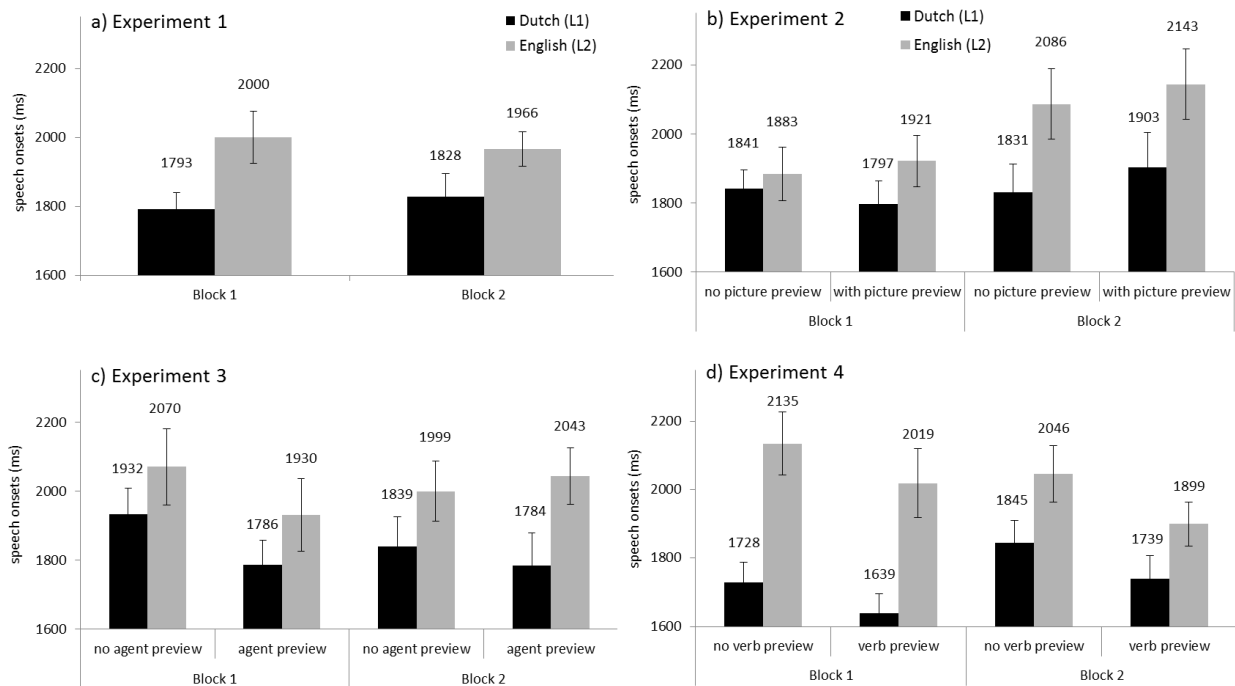


Figure 2. Speech onsets for active sentences (by-participant means with standard errors) produced as a function of (a) Language and Block in Experiment 1, and (b-d) Language, Block and Preview in Experiments 2-4. Error bars are standard errors.

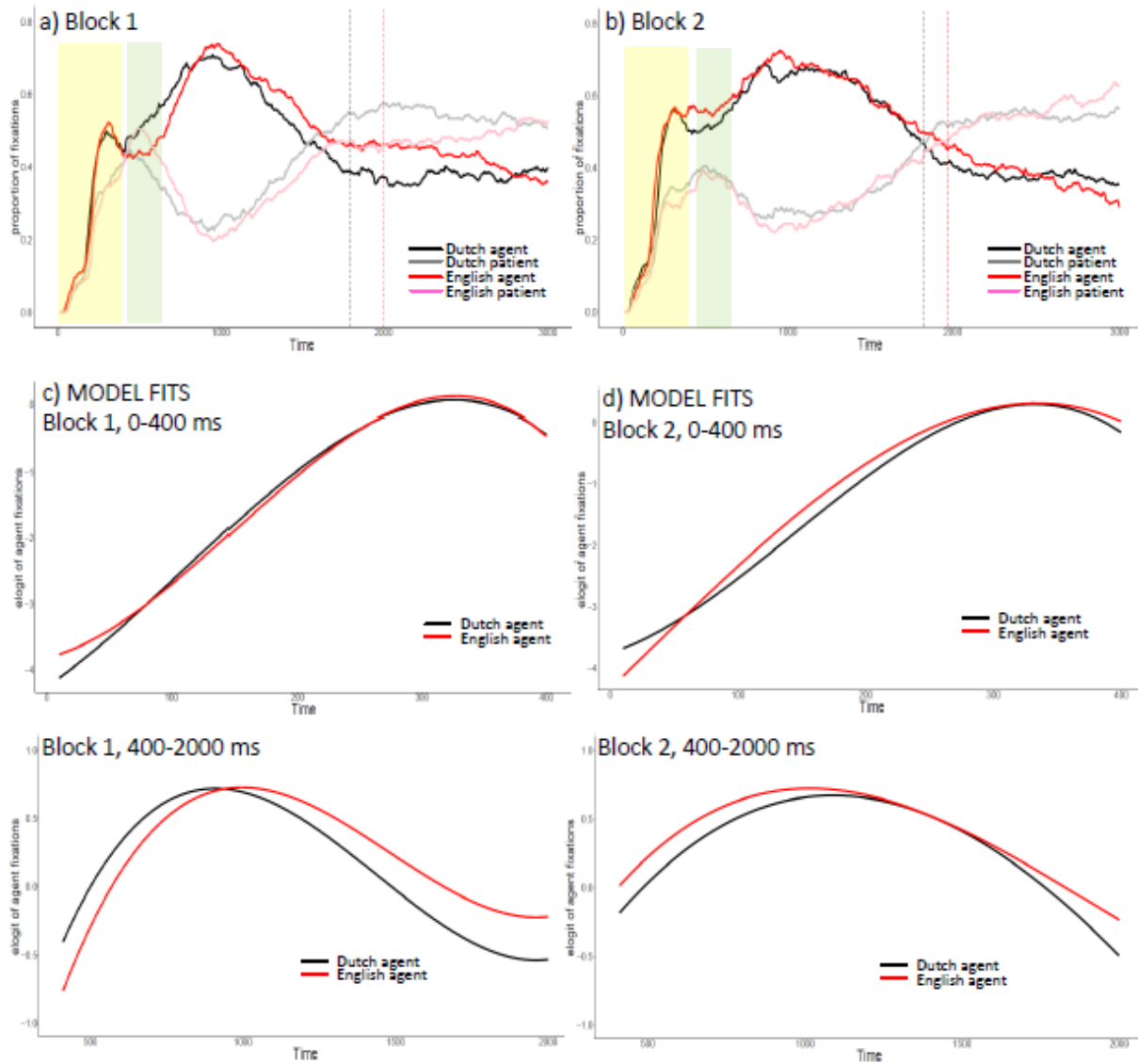


Figure 3. Time-course of formulation in Block 1 and Block 2 of Experiment 1: (a-b) proportions of by-participant agent-directed and patient-directed fixations and (c-d) model fits from by-participant analyses plotting agent-directed fixations (empirical logits; fits from the 0-400 ms are shown on a finer time scale for clarity). Patient fixations are complementary and thus are not plotted. The 0-400 ms time window is highlighted in yellow in all graphs; the 400-600 ms time window is highlighted in green for clarity in Figure (a) and (b). Vertical lines are speech onsets.

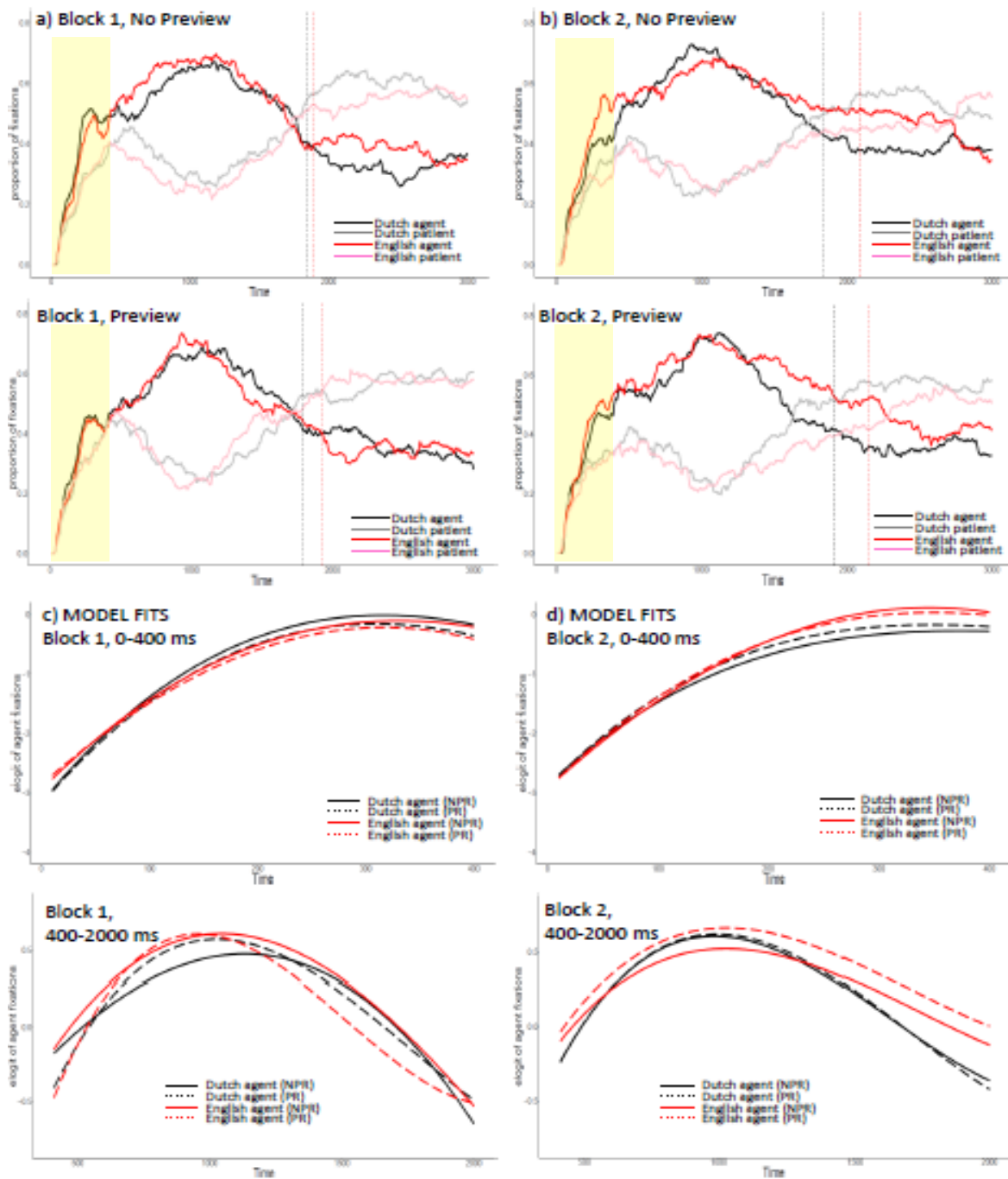


Figure 4. Time-course of formulation in Block 1 and Block 2 of Experiment 2: (a-b) by-participant agent-directed and patient-directed fixations (proportions), and (c-d) model fits from by-participant analyses. Vertical lines are speech onsets. PR=Preview, NPR=No Preview.

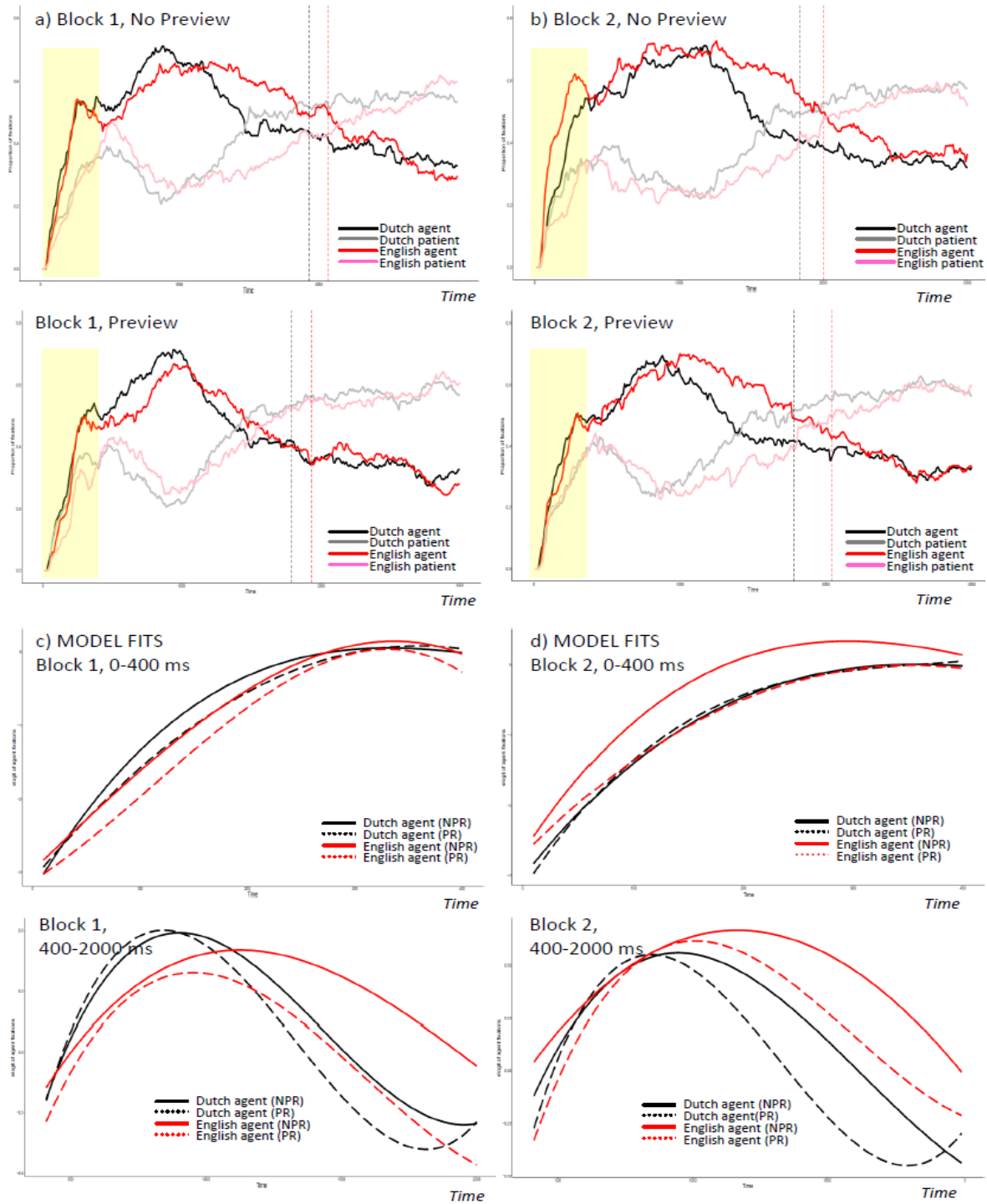


Figure 5. Time-course of formulation in Block 1 and Block 2 of Experiment 3: (a-b) by-participant agent-directed and patient-directed fixations, and (c-d) model fits from by-participant analyses (empirical logits). Vertical lines are speech onsets. PR=Preview, NPR=No Preview.

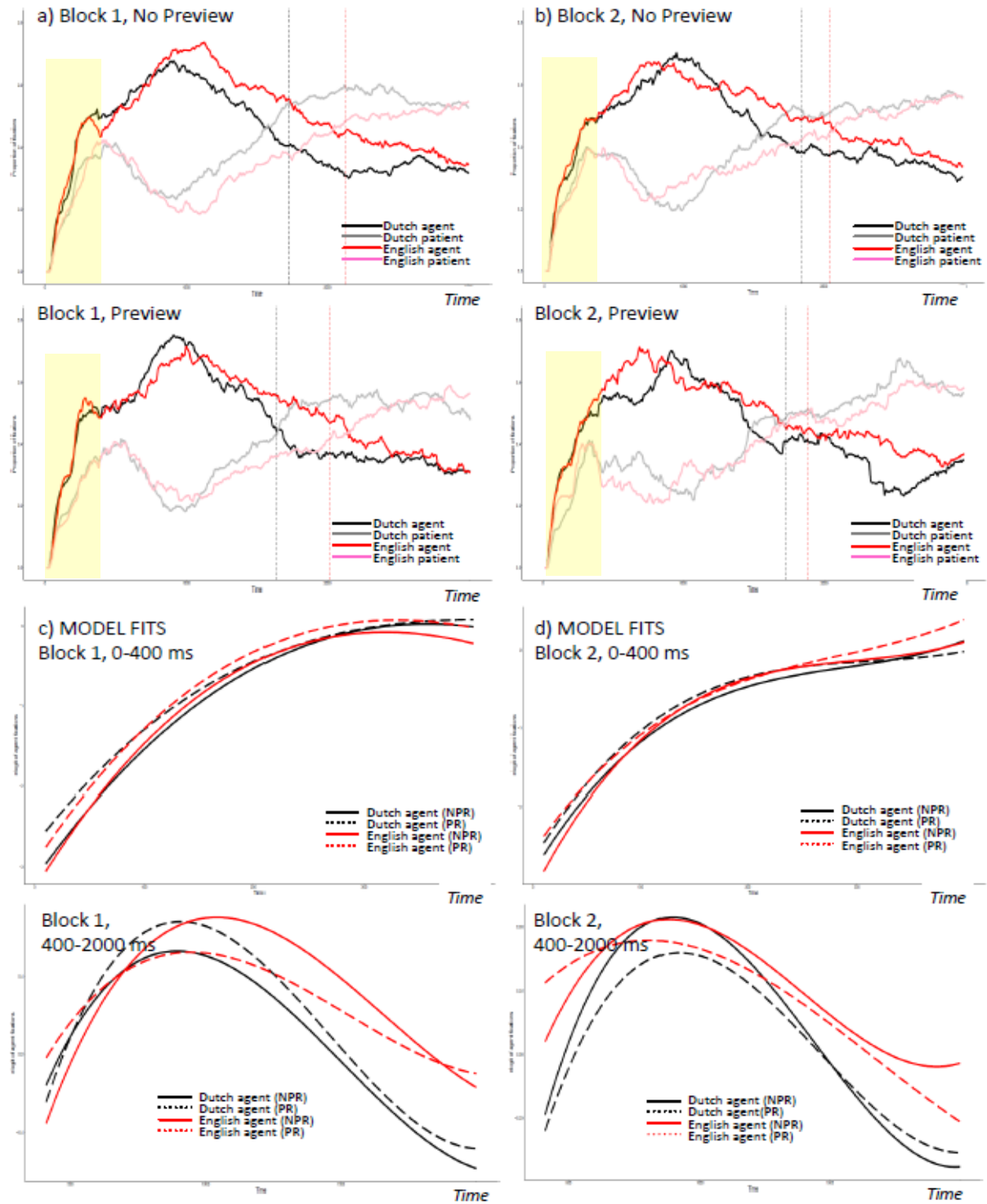


Figure 6. Time-course of formulation in Block 1 and Block 2 of Experiment 4: (a-b) by-participant agent-directed and patient-directed fixations, and (c-d) model fits from by-participant analyses (empirical logits). Vertical lines are speech onsets. PR=Preview, NPR=No Preview.

Appendix A

Target items used in Experiment 1-4, with the modal agent names, verbs, and patient names listed in parentheses in Dutch.

boxer punching cheerleader	(<i>bokser, slaan, cheerleader</i>)
dog chasing mailman	(<i>hond, achterna zitten, postbode</i>)
Santa Claus dragging Christmas tree	(<i>kerstman, slepen, kerstboom</i>)
girl taking cookie	(<i>meisje, pakken, koekje</i>)
man chopping log	(<i>man, hakken, boom</i>)
woman massaging man	(<i>vrouw, masseren, man</i>)
cameraman filming model	(<i>cameraman, filmen, model</i>)
man pulling donkey	(<i>man, meetrekken, ezel</i>)
girl tickling boy	(<i>meisje, kietelen, jongen</i>)
journalist interviewing actor	(<i>journaliste, interviewen, acteur</i>)
doctor washing baby	(<i>dokter, wassen, baby</i>)
boy letting birds out	(<i>jongen, vrijlaten, vogel</i>)
man pushing bed	(<i>man, duwen, bed</i>)
crab bit boy's foot	(<i>krab, bijten, jongen</i>)
mother dressing a boy	(<i>moeder, aankleden, jongen</i>)
bomb hitting ship	(<i>bom, raken, schip</i>)
boy kicking football	(<i>jongen, schoppen, voetbal</i>)
policeman stopping truck	(<i>politieagent, stoppen, vrachtwagen</i>)
pope crowning king/prince	(<i>paus, kronen, prins</i>)
dog guiding blind man	(<i>hond, leiden, blinde man</i>)
maid eating chocolate	(<i>dienstmeisje, eten, chocola</i>)
boyscout roasting pig	(<i>padvinder, roosteren, varken</i>)
cowboy caught sheriff	(<i>cowboy, vangen, sheriff</i>)
nurse holding boy	(<i>verpleegster, vasthouden, jongen</i>)
owl carrying/taking briefcase	(<i>uil, pakken, aktetas</i>)
bulldozer destroying building	(<i>hijskraan, slopen, gebouw</i>)
truck towing motorbike	(<i>sleepwagen, wegslepen, motor</i>)
cat scratching sofa	(<i>kat, kapotkrabben, bank</i>)
helicopter saving diver	(<i>helikopter, redden, duiker</i>)
man painting door	(<i>man, schilderen, deur</i>)
bride cutting cake	(<i>bruid, aansnijden, taart</i>)
frog catching fly	(<i>kikker, vangen, vlieg</i>)
man lifting armchair	(<i>man, optillen, stoel</i>)
windmill blowing farmer away	(<i>windmolen, wegblazen, boer</i>)
girl pushing boy on sled	(<i>meisje, duwen, jongen</i>)
thief stealing painting	(<i>dief, stelen, schilderij</i>)
punk spraying fence	(<i>punker, spuiten, hek</i>)
ambulance hitting woman	(<i>ambulance, aanrijden, vrouw</i>)
man revealing vase	(<i>man, onthullen, vaas</i>)
man reviving girl	(<i>man, reanimeren, meisje</i>)
robot destroying computer	(<i>robot, kapotmaken, computer</i>)
waiter kicking chef out	(<i>ober, trappen, kok</i>)
cat catching mouse	(<i>kat, pakken, muis</i>)
professor congratulating student	(<i>professor, feliciteren, studente</i>)
pirate burying treasure	(<i>piraat, begraven, schat</i>)
ballerina slapping pianist	(<i>ballerina, slaan, pianist</i>)
man planting tree	(<i>man, planten, boom</i>)
lightning striking church	(<i>bliksem, raken, kerk</i>)
barber shaving prisoner	(<i>kapper, scheren, gevangene</i>)
pelican puncturing hot-air balloon	(<i>pelikaan, lekprikken, luchtballon</i>)
girl throwing away present	(<i>meisje, weggooien, cadeau</i>)
man brushing suit	(<i>man, borstelen, pak</i>)
burglar smashing window	(<i>inbreker, kapotslaan, raam</i>)
bee stinging man	(<i>bij, steken, man</i>)

man lifting bench	(<i>man, optillen, bank</i>)
train crushing bus	(<i>trein, door midden rijden, bus</i>)
horse kicking cow	(<i>paard, schoppen, koe</i>)
fisherman catching fish	(<i>visser, vangen, vis</i>)
man threatening woman	(<i>man, bedreigen, vrouw</i>)
leopard catching snake	(<i>luipaard, vangen, slang</i>)
paparazzi photographing queen	(<i>paparazzi, fotograferen, koningin</i>)
detective finding gun	(<i>detectif, vinden, pistool</i>)
dog catching butterfly	(<i>hond, vangen, vlinder</i>)
man fixing car	(<i>man, repareren, auto</i>)
Inuit chasing away bear	(<i>Eskimo, wegjagen, ijsbeer</i>)
girl pulling suitcase	(<i>meisje, trekken, koffer</i>)
fireman saving child	(<i>brandweerman, redden, jongen</i>)
old woman cleaning stairs	(<i>oude vrouw, dweilen, trap</i>)
grandfather kissing baby	(<i>opa, kussen, baby</i>)
tiger scratching man	(<i>tijger, krabben, man</i>)
alien taking astronaut with him	(<i>alien, meenemen, astronaut</i>)
girl hanging sheet	(<i>meisje, ophangen, laken</i>)
shark eating man	(<i>haai, eten, man</i>)
policeman arresting man	(<i>politieagent, arresteren, man</i>)
stork bringing baby	(<i>ooievaar, brengen, baby</i>)
swimmer pushing paparazzi	(<i>zwemmer, duwen, paparazzi</i>)
bodyguard protecting president	(<i>bodyguard, beschermen, president</i>)
army attacking castle	(<i>leger, aanvallen, kasteel</i>)
elephant lifting clown	(<i>olifant, optillen, clown</i>)
man making statue	(<i>man, maken, standbeeld</i>)

Supplementary Material A.

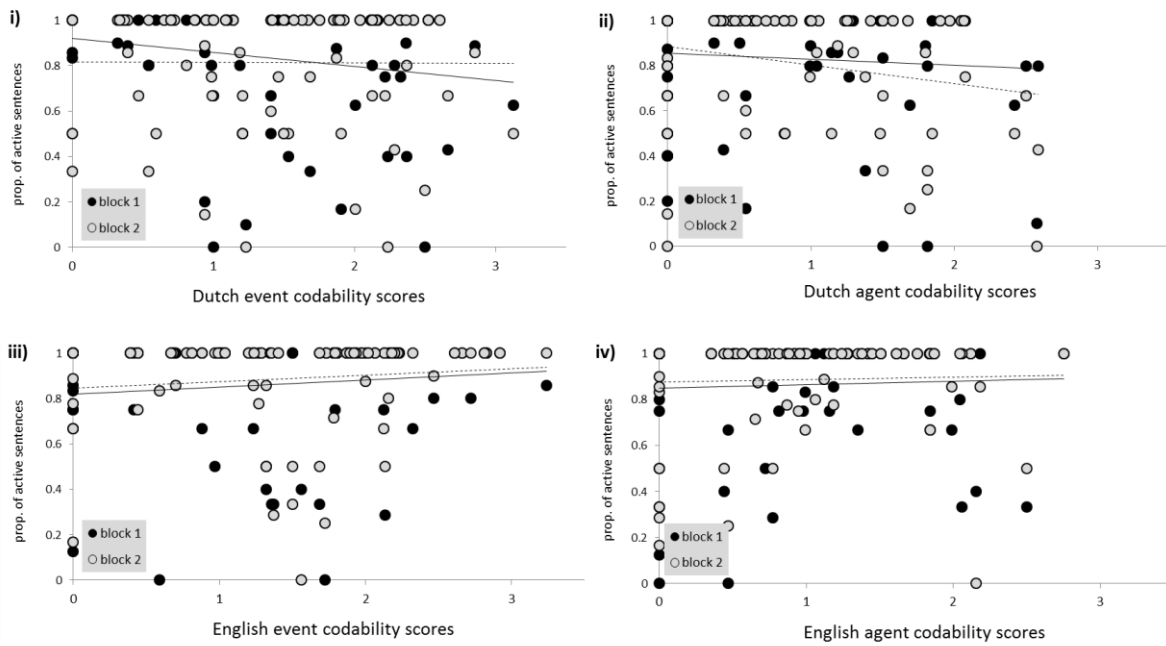
Counts of active and passive sentences produced across experiments (truncated passives were excluded from analysis).

	actives	passives
Experiment 1		
Dutch	805	146
English	827	100
Experiment 2		
Dutch: <i>preview, no preview</i>	506, 506	85, 74
English: <i>preview, no preview</i>	509, 470	66, 52
Experiment 3		
Dutch: <i>preview, no preview</i>	425, 403	38, 70
English: <i>preview, no preview</i>	378, 343	36, 49
Experiment 4		
Dutch: <i>preview, no preview</i>	398, 653	67, 94
English: <i>preview, no preview</i>	416, 623	54, 79

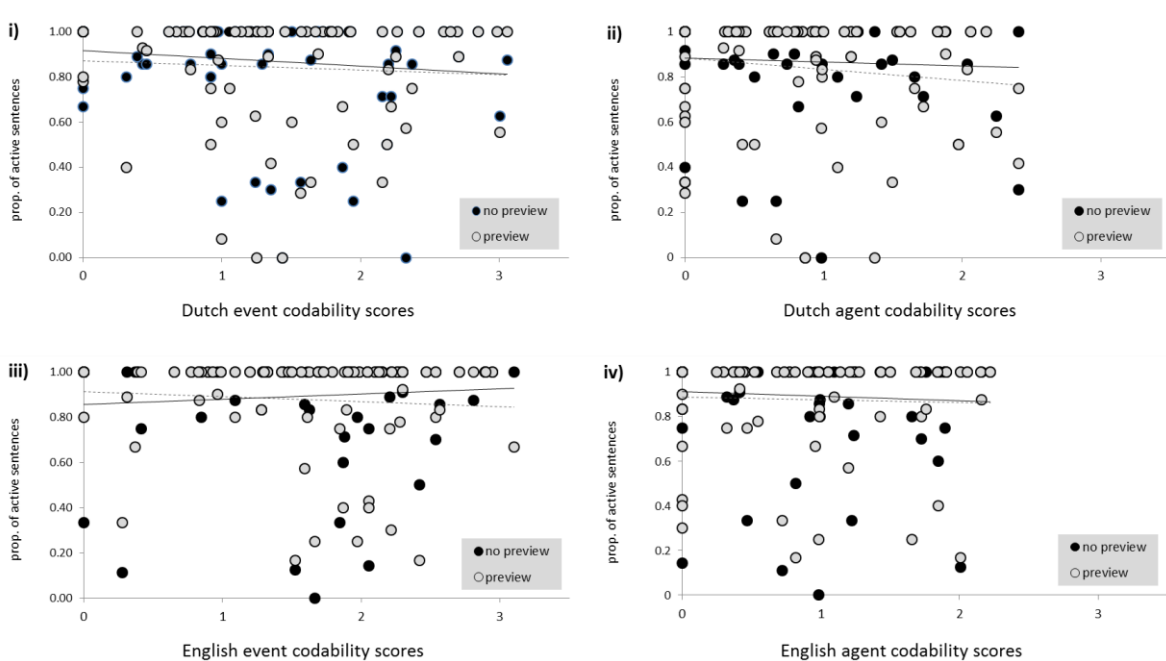
Supplementary Material B.

Proportion of active sentences (by-item means) relative to Event and Agent Codability in (a-d) Experiments 1-4. Panels (i) and (ii) for each experiment show results for Dutch Event and Agent codability; panels (iii) and (iv) show results for English Event and Agent codability. Items with lower codability scores are high-codability items (i.e., items with little variability in responses) and items with codability scores of 0 are items with perfect name agreement. Items with higher codability scores are low-codability items (i.e., items with high variability in responses).

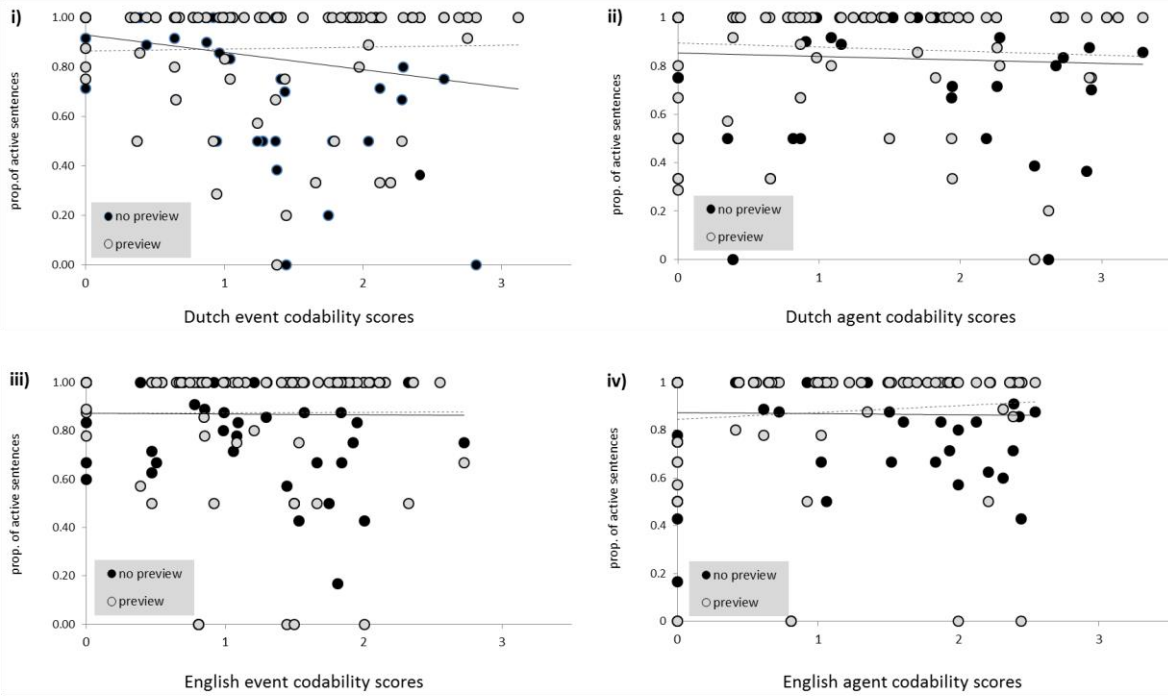
a) Experiment 1



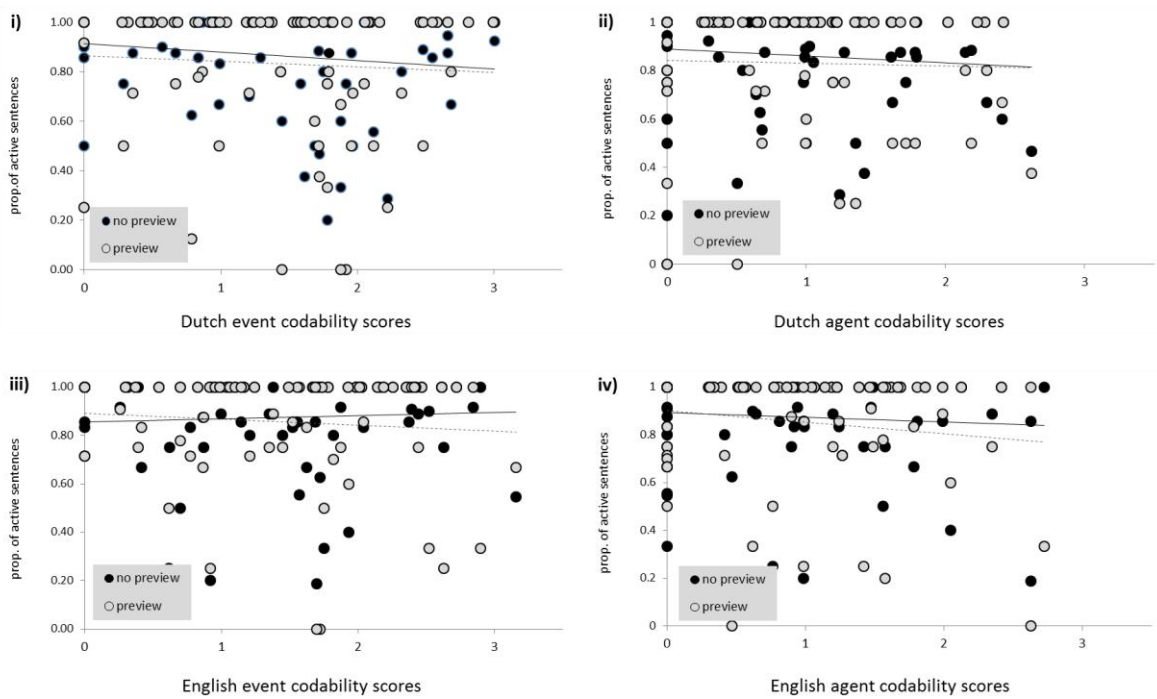
b) Experiment 2



c) Experiment 3



d) Experiment 4



Supplementary Material C.

GCA results for the joint analysis of Blocks 1 and 2 in Experiment 1. All models included random by-participant or by-item slopes for Language and Block.

Effect	By-participants			By-items		
	Est.	SE	t	Est.	SE	t
<i>(a) 0-400 ms</i>						
Intercept	-1.25	.07	-17.13*	-.88	.08	-10.71*
Linear	8.16	.34	24.18*	5.45	.29	18.91*
Quadratic	-2.49	.22	-11.16*	-.96	.19	-5.01*
Cubic	-.96	.21	-4.57*	-1.19	.11	-10.56*
Language	.03	.07	.44	-.03	.05	-.57
Block	.25	.07	3.42*	.14	.05	2.66*
Linear * Language	.09	.14	.64	-.02	.08	-.31
Quadratic * Language	-.16	.13	-1.31	-.02	.08	-.22
Cubic * Language	.15	.11	1.39	.14	.07	1.88†
Linear * Block	-.12	.14	-.84	-.03	.08	-.35
Quadratic * Block	-.12	.13	-.94	-.04	.08	-.48
Cubic * Block	.48	.11	4.57*	.25	.07	3.27*
Language * Block	.06	.29	.22	-.04	.02	-1.81†
Linear * Language * Block	.34	1.35	.26	.18	.16	1.15
Quadr. * Language * Block	-1.13	.89	-1.27	-.53	.15	-3.50*
Cubic * Language * Block	1.10	.84	1.31	.23	.15	1.53
<i>(b) 400-2000 ms</i>						
Intercept	.28	.04	7.85*	.26	.05	4.72*
Linear	-1.95	.35	-5.61*	-1.05	.42	-2.47*
Quadratic	-3.59	.24	-15.25*	-2.82	.39	-7.32*
Cubic	1.19	.25	4.58*	.89	.30	2.95*
Language	.08	.07	1.22	.11	.04	2.94*
Block	.15	.07	2.26*	.07	.03	2.16*
Linear * Language	1.36	.07	18.15*	.87	.08	11.57*
Quadratic * Language	.15	.08	1.97*	.08	.08	1.12
Cubic * Language	-.07	.08	-.93	.13	.08	1.67
Linear * Block	.79	.07	10.56*	.42	.08	5.46*
Quadratic * Block	.47	.08	6.28*	.45	.08	5.89*
Cubic * Block	-1.73	.08	-23.00*	-1.27	.08	-16.70*
Language * Block	.00	.14	.03	.02	.01	1.55
Linear * Language * Block	-2.46	1.39	-1.78†	-2.32	.15	-13.37*
Quadr. * Language * Block	1.51	.94	1.60	1.54	.15	10.20*
Cubic * Language * Block	.13	1.04	.12	.27	.15	1.81†

* $p < .05$ † $p < .10$

Supplementary Material D.

GCA results for the joint analysis of Blocks 1 and 2 in Experiment 2 (with all three-way interactions). The intercept and Time terms are omitted in the Table. All models included by-participant or by-item random slopes for Language, Preview, and Block.

Effect	By-participants			By-items		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
<i>0-400 ms</i>						
Language	.06	.07	.77	.04	.05	.86
Block	.03	.07	.44	.04	.06	.74
Preview	-.01	.08	-.17	-.03	.04	-.94
Language * Block	.17	.15	1.09	.13	.02	6.88*
Language * Preview	-.04	.02	-1.76	.001	.02	.08
Block * Preview	.18	.02	7.99*	.09	.02	4.67*
Language * Block * Preview	-.11	.31	-.36	-.11	.04	-3.01*
Linear * Language	.19	.08	2.37*	.07	.06	1.15
Quadratic * Language	.14	.07	1.88†	.09	.06	1.42
Cubic * Language	-.14	.07	-1.99*	-.07	.06	-1.20
Linear * Block	.10	.08	1.25	.19	.06	3.04*
Quadratic * Block	.38	.08	4.96*	.28	.06	4.70*
Cubic * Block	-.00	.08	-.01	.07	.06	1.17
Linear * Preview	-.02	.08	-.25	-.04	.06	-.67
Quadratic * Preview	-.03	.08	-.45	-.05	.06	-.86
Cubic * Preview	-.06	.07	-.80	.04	.06	.67
Linear * Lang * Block	1.18	.63	1.87†	.98	.12	8.11*
Quadratic * Lang * Block	-.37	.51	-.71	-.21	.12	-1.74
Cubic * Lang * Block	-.11	.38	-.28	-.07	.11	-.63
Linear * Lang * Preview	-.31	.16	-1.96*	-.16	.12	-1.33
Quadratic * Lang * Preview	.10	.15	.64	-.07	.12	-.55
Cubic * Lang. * Preview	.11	.14	.75	.16	.12	1.37
Linear * Block * Preview	.39	.16	2.49*	.11	.12	.93
Quadratic * Block * Preview	-.12	.15	-.79	.09	.12	.77
Cubic * Block * Preview	.15	.14	1.05	.11	.12	.94
<i>400-2000 ms</i>						
Language	.05	.04	1.20	.01	.03	.50
Block	.09	.04	1.98*	.12	.04	2.77*
Preview	.00	.03	.04	.01	.03	.49
Language * Block	.09	.11	.83	.10	.01	9.35*
Language * Preview	-.00	.01	-.20	-.05	.01	-4.54*
Block * Preview	.12	.01	10.43*	.03	.01	2.74*
Lang * Block * Preview	.29	.12	2.46*	.10	.02	4.56*
Linear * Language	.07	.07	.96	.16	.06	2.45*
Quadratic * Language	.49	.07	6.93*	.22	.06	3.40*
Cubic * Language	.12	.07	1.71	.07	.06	1.11
Linear * Block	.30	.07	4.26*	.04	.06	.61
Quadratic * Block	.83	.07	11.73*	.69	.06	10.75*

Cubic * Block	-.06	.07	-.78	-.05	.06	-.85
Linear * Preview	-.21	.07	-3.00*	-.03	.06	-.44
Quadratic * Preview	-.12	.07	-1.79†	-.12	.06	-1.88
Cubic * Preview	.46	.07	6.57*	.43	.06	6.77*
Linear * Lang * Block	1.70	1.30	1.31	1.45	.13	11.40*
Quadratic * Lang * Block	.98	.83	1.18	.63	.13	4.94*
Cubic * Lang * Block	-.90	.64	-1.40	-1.17	.13	-9.23*
Linear * Lang * Preview	-.37	.14	-2.70*	-.38	.13	-3.01*
Quadratic * Lang * Preview	-.11	.14	-.77	.04	.13	.35
Cubic * Lang * Preview	.21	.14	1.55	.27	.13	2.15*
Linear * Block * Preview	.46	.14	3.32*	.42	.13	3.33*
Quadratic * Block * Preview	.14	.14	1.03	.36	.13	2.86*
Cubic * Block * Preview	-1.04	.12	-7.51*	-.97	.13	-7.78*

* $p < .05$ † $p < .10$

Supplementary Material E.

GCA results for the joint analysis of Blocks 1 and 2 in Experiment 3 (with all three-way interactions). The intercept and Time terms are omitted in the Table. All models included by-participant or by-item slopes for Language, Preview, and Block.

Effect	By-participants			By-items		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
<i>0-400 ms</i>						
Language	.04	.06	.66	-.04	.04	-.90
Block	.18	.06	2.74*	.13	.04	3.02*
Preview	-.19	.06	-3.17*	-.05	.05	-.97
Language * Block	.38	.16	2.34*	.23	.02	10.92*
Language * Preview	-.20	.03	-7.54*	-.19	.02	-8.70*
Block * Preview	-.08	.03	-2.91*	.03	.02	1.39
Language * Block * Preview	-.26	.24	-1.07	.02	.04	.37
Linear * Language	.05	.09	.58	-.19	.07	-2.97*
Quadratic * Language	.14	.09	1.58	.10	.06	1.52
Cubic * Language	-.37	.08	-4.63*	-.23	.06	-3.58*
Linear * Block	-.81	.09	-8.93*	-.43	.06	-6.64*
Quadratic * Block	-.07	.09	-.75	-.24	.06	-3.68*
Cubic * Block	.35	.08	4.29*	.26	.06	4.21*
Linear * Preview	.26	.09	2.89*	-.17	.07	-2.46*
Quadratic * Preview	.26	.09	3.01*	-.03	.07	-.51
Cubic * Preview	-.20	.08	-2.49*	-.02	.07	-.34
Linear * Lang * Block	-.80	.76	-1.04	-.62	.13	-4.84*
Quadratic * Lang * Block	-.29	.49	-.60	-.50	.13	-3.92*
Cubic * Lang * Block	.45	.54	.83	.30	.13	2.37*
Linear * Lang * Preview	.28	.18	1.53	.24	.13	1.82†
Quadratic * Lang * Preview	.15	.17	.86	.25	.13	1.91*
Cubic * Lang. * Preview	-.34	.16	-2.10*	-.15	.13	-1.19
Linear * Block * Preview	-.25	.18	-1.37	-.37	.13	-2.88*
Quadratic * Block * Preview	.08	.17	.46	-.10	.13	-.79
Cubic * Block * Preview	.28	.16	1.75†	.38	.13	3.04*
<i>400-2000 ms</i>						
Language	.14	.05	2.69*	.14	.03	4.10*
Block	.11	.05	2.34*	.07	.03	2.00*
Preview	-.17	.04	-4.09*	-.11	.04	-3.01*
Language * Block	.19	.12	1.51	.20	.01	17.20*
Language * Preview	-.07	.01	-5.90*	-.05	.01	-3.90*
Block * Preview	.03	.01	2.73*	.03	.01	2.15*
Language * Block * Preview	.12	.16	.76	.00	.02	.03
Linear * Language	2.19	.08	28.59*	1.77	.07	25.18*
Quadratic * Language	-.11	.08	-1.37	-.03	.07	-.48
Cubic * Language	-1.16	.08	-15.09*	-.63	.07	-8.95*
Linear * Block	.52	.08	6.72*	.39	.07	5.62*
Quadratic * Block	-.03	.08	-.34	.05	.07	.73

Cubic * Block	-.07	.08	-.92	-.07	.07	-.99
Linear * Preview	-1.28	.08	-16.75*	-.69	.08	-9.13*
Quadratic * Preview	.18	.08	2.42*	-.03	.07	-.37
Cubic * Preview	.90	.08	11.83*	.75	.07	9.96*
Linear * Lang * Block	.22	1.32	.17	.53	.14	3.78*
Quadratic * Lang * Block	.23	.84	.28	.12	.14	.87
Cubic * Lang * Block	.30	.86	.35	.56	.14	4.03*
Linear * Lang * Preview	-.21	.15	-1.35	-.11	.14	-.78
Quadratic * Lang * Prev.	-1.02	.15	-6.65*	-.98	.14	-6.98*
Cubic * Lang * Preview	-.34	.15	-2.21*	-.45	.14	-3.18*
Linear * Block * Preview	.53	.15	3.47*	.52	.14	3.69*
Quadratic * Block * Preview	.20	.15	1.31	.24	.14	1.71
Cubic * Block * Preview	.29	.15	1.89†	.44	.14	3.13*

* $p < .05$ † $p > .10$

Supplementary Material F.

GCA results for the joint analysis of Blocks 1 and 2 in Experiment 4 (with all three-way interactions). The intercept and Time terms are omitted in the Table. All models included by-participant or by-item random slopes for Language, Preview, and Block, except model (b) which included by-participant random slopes for Preview and Block.

Effect	By-participants			By-items		
	Est.	SE	<i>t</i>	Est.	SE	<i>T</i>
<i>(a) 0-400 ms</i>						
Language	.03	.07	.43	-.03	.04	-.66
Block	.09	.07	1.32	.08	.05	1.66
Preview	-.14	.06	-2.24*	.08	.05	1.70
Language * Block	.02	.16	.10	-.06	.02	-3.32*
Language * Preview	.05	.02	2.34*	-.03	.02	-1.61
Block * Preview	.09	.03	3.45*	-.06	.02	-3.02*
Language * Block * Preview	.03	.24	.13	-.06	.04	-1.62
Linear * Language	.27	.08	3.35*	.08	.06	1.33
Quadratic * Language	-.20	.08	-2.56*	-.20	.06	-3.30*
Cubic * Language	-.03	.07	-.40	-.01	.06	-.22
Linear * Block	-.73	.08	-9.17*	-.59	.06	-9.70*
Quadratic * Block	.33	.08	4.20*	.05	.06	.75
Cubic * Block	.45	.07	6.27*	.58	.06	10.00*
Linear * Preview	-.39	.08	-4.89*	-.43	.06	-7.02*
Quadratic * Preview	.26	.08	3.30*	.26	.06	4.27*
Cubic * Preview	-.03	.07	-.39	-.01	.06	-.16
Linear * Lang * Block	.59	.67	.87	.57	.12	4.80*
Quadratic * Lang * Block	.48	.45	1.05	.45	.12	3.85*
Cubic * Lang * Block	-.17	.44	-.39	-.04	.11	-.37
Linear * Lang * Preview	.57	.16	3.66*	.40	.12	3.37*
Quadratic * Lang * Prev.	.55	.15	3.52*	.38	.12	3.18*
Cubic * Lang. * Preview	-.20	.14	-1.38	-.05	.12	-.41
Linear * Block * Preview	.21	.16	1.33	.35	.12	2.89*
Quadratic * Block * Preview	.08	.15	.51	.13	.12	1.07
Cubic * Block * Preview	-.10	.14	-.72	-.17	.12	-1.47
<i>(b) 400-2000 ms</i>						
Language	.15	.04	3.56*	.08	.03	2.61*
Block	-.05	.04	-1.25	-.05	.03	-1.91†
Preview	-.02	.04	-.58	-.02	.03	-.68
Language * Block	.01	.10	.08	.04	.01	4.05*
Language * Preview	-.05	.01	-4.55*	-.07	.01	-6.74*
Block * Preview	-.06	.01	-5.36*	-.01	.01	-.48
Language * Block * Preview	.16	.15	1.10	.14	.02	6.37*
Linear * Language	1.13	.07	15.84*	1.11	.07	17.04*
Quadratic * Language	.78	.07	10.97*	.63	.07	9.62*
Cubic * Language	-.60	.07	-8.46*	-.40	.06	-6.23*

Linear * Block	.03	.07	.36	-.02	.07	-.37
Quadratic * Block	.96	.07	13.37*	.83	.07	12.78*
Cubic * Block	-.20	.07	-2.85*	-.05	.06	-.73
Linear * Preview	-.06	.08	-.79	-.27	.07	-4.10*
Quadratic * Preview	.50	.07	7.06*	.40	.07	6.11*
Cubic * Preview	-.18	.07	-2.49*	.06	.07	.86
Linear * Lang * Block	-1.67	1.31	-1.28	-.76	.13	-5.97*
Quadratic * Lang * Block	1.14	.80	1.42	1.09	.13	8.50*
Cubic * Lang * Block	.08	.74	.10	-.33	.13	-2.56*
Linear * Lang * Preview	-1.07	.14	-7.53*	-1.15	.13	-8.79*
Quadratic * Lang * Prev.	1.15	.14	8.13*	.94	.13	7.20*
Cubic * Lang * Preview	-.45	.14	-3.18*	.04	.13	.29
Linear * Block * Preview	.30	.14	2.13*	.32	.13	2.44*
Quadratic * Block * Preview	-.17	.14	-1.19	.20	.13	1.51
Cubic * Block * Preview	-.67	.14	-4.71*	-.79	.13	-6.07*

* $p < .05$ † $p < .10$